Exhibit 111

Conditions Associated with Antibodies Against the Tumor-Associated Antigen MUC1 and Their Relationship to Risk for Ovarian Cancer

Daniel W. Cramer,[1] Linda Titus-Ernstoff,[4] John R. McKolanis,[5] William R. Welch,[2] Allison F. Vitonis,[1] Ross S. Berkowitz,[3] and Olivera J. Finn[5]

[1]Ob-Gyn Epidemiology Center, Departments of [2]Pathology (Women's and Perinatal Pathology Division) and [3]Obstetrics and Gynecology, Division of Gynecologic Oncology, Brigham and Women's Hospital, Boston, Massachusetts; [4]Norris Cotton Cancer Center, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire; and [5]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

Abstract

Many cancers, including ovarian, overexpress epithelial mucin (MUC1) and promote anti-MUC1 antibodies that may correlate with more favorable prognosis. By extension, risk for ovarian cancer might be reduced by preexisting MUC1-specific immunity. We measured anti-MUC1 antibodies in 705 control women, identified events predicting antibodies, and estimated ovarian cancer risk by comparing profiles of events generating antibodies in controls with those in 668 ovarian cancer cases. Factors predicting antibodies included oral contraceptive use, breast mastitis, bone fracture or osteoporosis, pelvic surgeries, nonuse of talc in genital hygiene, and to a lesser extent intrauterine device use and current smoking. There was a significant increase in the likelihood of having anti-MUC1 antibodies from 24.2% in women with 0 or 1 condition, to 51.4% in those with five or more conditions. By the same index of events, the risk for ovarian cancer was inversely associated with number of conditions predisposing to anti-MUC1 antibodies. Compared with having experienced 0 or 1 event, the adjusted risk for ovarian cancer decreased progressively with relative risks (and 95% confidence limits) of 0.69 (0.52-0.92), 0.64 (0.47-0.88), 0.49 (0.34-0.72), and 0.31 (0.16-0.61), respectively for women with two, three, four, and five or more events related to the presence of antibodies (P_{trend} < 0.0001.) We conclude that several traditional and new risk factors for ovarian cancer may be explained by their ability to induce MUC1 immunity through exposure of MUC1 to immune recognition in the context of inflammatory or hormonal processes in various MUC1-positive tissues. (Cancer Epidemiol Biomarkers Prev 2005;14(5):1125 31)

Introduction

Human mucin (MUC) family member, MUC1, is a high molecular weight protein expressed in a highly glycosylated form and low levels by many types of normal epithelial cells and in a hypoglycosylated form and high levels by most epithelial adenocarcinomas, including breast and ovarian cancer (1). Anti-MUC1 antibodies have been described and correlated with a more favorable prognosis (2-5) showing that patients generate immunity against MUC1 produced by their tumors and defining MUC1 as a tumor-associated antigen and candidate for cancer vaccines (6). Anti-MUC1 antibodies are also found in healthy individuals, especially in women during pregnancy and lactation. It has been hypothesized that a natural immunity against tumor MUC1 might develop and account for the long-term protective effect of pregnancy or breast-feeding on breast cancer risk (7), an elaboration on the so called ''fetal antigen theory'' (8). Indeed it has been shown that sera from multiparous women, but not from nulliparous women or from men, are able to mediate killing of breast cancer cells (9). Supporting a key role for MUC1 in these reactions, core peptide sequences from MUC1 can induce proliferation of T cells and cytotoxic T-cell responses in multiparous women (10). Recently, the ''fetal antigen'' hypothesis was extended to ovarian cancer after it was shown that sera from multiparous women also reacted with multiple antigens from ovarian cancer cells more strongly than sera from nulliparous women or men (11), although MUC1 was not specifically examined in these experiments.

In this study, we used an ELISA to determine the presence and relative amounts of MUC1-specific antibody in women from the general population who served as controls in a study of ovarian cancer. In analyses confined to these controls, we identified the predictors of anti-MUC1 antibodies and used case-control comparisons to evaluate these predictors in relation to ovarian cancer risk. We hypothesized that events predicting anti-MUC1 antibodies would be inversely associated with ovarian cancer risk and that there would be a cumulative effect of such events.

Materials and Methods

Subject Recruitment and Study. This report is based on the second phase of a population-based case-control study of ovarian cancer conducted between 7/98 and 7/03 and involving eastern Massachusetts and all of New Hampshire, approved by the Brigham and Women's Hospital and Dartmouth Medical Center's Institutional Review Boards. We identified 1,267 cases from tumor boards and Statewide Registries and excluded 119 cases who died, 110 who moved from the study area, one who had no telephone, 23 who did not speak English, and 46 found to have a nonovarian primary upon review. Of the remaining 968, physicians denied permission to contact 106 and 171 declined to participate, leaving 691 cases interviewed. Of these, 668 had an epithelial ovarian cancer (including borderline malignancies) and are included in this report. A small number of cases (n = 48) were enrolled before surgery.

Received 1/17/05; accepted 3/4/05.

Grant support: National Cancer Institute grants CA 054419 (D.W. Cramer), CA 86381 (D.W. Cramer), and CA 56103 (O.J. Finn).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Daniel W. Cramer, Epidemiology Center, Brigham and Women's Hospital, 221 Longwood Avenue, Boston, MA 02115. Phone: 617-732-4895; Fax: 617-732-4899. E-mail: dcramer@partners.org

Controls were identified through town books in Massachusetts and Drivers' License lists in New Hampshire and sampled to match the age and residence of previously accumulated cases. Invitations to participate were sent to 1843 potential controls. Of these 318 had moved and could not be located or had died, 197 (in Massachusetts) could not be recontacted because subjects returned an "opt out" postcard required by the hospital's Institutional Review Boards, and 47 no longer had a working telephone. Of the remaining 1,281 who were contacted, 152 were ineligible because they had no ovaries or were not the correct age, 59 were incapacitated or did not speak English, and 349 declined, leaving 721 who were interviewed and included in this report.

After written informed consent, an in-person interview dealing with demographic, medical, and family history was conducted. Subjects also completed a self-administered dietary questionnaire. Heparinized blood specimens were collected from subjects agreeing to provide one; separated into red cell, buffy coat, and plasma components; and stored at 80°C.

ELISA Assay for Anti-MUC1 Antibodies. Plasma specimens were available for measuring anti-MUC1 antibodies in 48 cases with preoperative bloods and 705 controls. Antibodies were measured against a synthetic 100-mer MUC1 peptide corresponding to five tandem repeats of the MUC1 polypeptide core tandem repeat region, according to our previously published protocol (2). Briefly, 0.5 μg of MUC1 peptide in 100 μL of PBS was added to each well of Immulon 4 plates (Dynax, Chantilly, VA) and incubated overnight at 4°C. Control plates without the MUC1 peptide were also prepared. The plates were washed thrice with PBS and 100 μL of 2.5% bovine serum albumin in PBS added for 1 hour at room temperature to coat remaining sites in the well (blocking step). Fifty microliters of serially diluted plasma (1:20 to 1:80 in PBS) were added to the MUC1 peptide-coated and control plates and incubated for 1 hour at room temperature. The plates were washed 5× with 100 μL PBS and 0.1% Tween 20 detergent. Fifty microliters of secondary antibody, alkaline phosphatase labeled goat anti-human polyvalent IgM, IgG, IgA (Sigma-Aldrich, St. Louis, MO), diluted 1:1,000, was added for 1 hour at room temperature, and plates again washed 5× with PBS-Tween. One hundred microliters of alkaline phosphatase substrate pNPP (Sigma-Aldrich) were added at 3 mg/mL in 0.05 mol/L $NaCO_3$ and 0.5 mmol/L $MgCl_2$ and the plates incubated in the dark for exactly 1 hour before adding 50 μL of the stop solution (0.5 mol/L NaOH). The absorbance at 405 to 410 nm was measured using the plate reader MRX Revelation (Thermo Labsystems, Chantilly, VA). Absorbance values for each sample in the MUC1-coated plate were compared with values in the antigen-negative plates to subtract nonspecific binding. Based upon the previous responses in over 500 cancer cases and controls, absorbance reactions at the 1:20 dilution at <0.6 are scored as negative, reactions in the 0.6 to 0.79 range as low, reactions in the 0.8 to 0.99 range as intermediate, and reactions ≥1.0 as high. In the current study, 20 blinded replicate specimens were included and the Spearman correlation coefficient between the paired absorbances was 0.93 ($P < 0.0001$).

Statistical Methods. Logistic regression analysis was used to compare those with an antibody reaction at any level against those considered negative for MUC1 antibody ($A < 0.6$), while adjusting for potential confounding variables. Spearman rank correlations or generalized linear modeling was used to assess differences in absorbance levels (using log-transformed values of absorbance) for a more quantitative assessment of factors affecting anti-MUC1 antibody production. Combinations of factors were examined to identify the best cumulative index of experiences associated with likelihood of having antibodies. Ovarian cancer cases and controls were then categorized by the presence or absence of events found to affect the likelihood of antibodies and risk for ovarian cancer calculated using unconditional multivariate logistic regression to adjust for potential confounders. In our models, we adjusted for the matching variables, age (continuous), and study site (Massachusetts, New Hampshire), as well as ethnicity (White, non-White), religious background (Jewish, non-Jewish), and parity as a continuous variable except where noted in the text or footnotes.

Results

The distributions of absorbance readings (corresponding to the amount of anti-MUC1 antibodies measured in the ELISA assay) seemed bimodal in cases with preoperative bloods and skewed right in controls prompting log transformation for statistical testing (Fig. 1). By a cutoff of $A \geq 0.6$, 33.8% of controls and 45.8% of cases were positive for antibodies. By a cutoff of $A \geq 1.0$, 12.3% of controls and 25% of cases had a high level of antibodies, a significant difference that likely reflects an ongoing immune response to tumor in the cases.

Events Predicting Occurrence of Anti-MUC1 Antibodies. A number of demographic, reproductive, and medical conditions were examined as they affected the likelihood of controls having a low, intermediate, high level, or any anti-MUC1 antibody (Table 1). The last two columns show the (geometric) mean absorbance value, its SE, and the P from the linear regression model. Age was a strong predictor with 50% having antibodies at ages <35, declining to 29.3% at ages 55 to 64 years, and increasing back up to 32.6% in those ages ≥65 years, prompting age adjustment when testing for the significance of further variables. The proportion of women who were positive for anti-MUC1 antibody was similar for women who had never been pregnant (33.3%), had at least one live birth (34.1%), or had breast-fed without experiencing a mastitis (33.0%) but elevated for women who had experienced mastitis while breast-feeding (46.1%). Notably, 25.0% of women reporting mastitis had high antibody levels compared with 10% to 14% of parous women who either never breast-fed or breast-fed and reported no mastitis ($P = 0.05$). Women who had used oral contraceptives (OC), compared with those who had not, were more likely to have antibodies; and this was most apparent among premenopausal women in whom 40.7% of OC users had antibodies compared with 26.7% of nonusers ($P = 0.05$). The proportion of women with antibodies was also higher for those who reported a bone fracture or osteoporosis after age 40 or within 20 years of their age at interview (36.0%) than in those who had not (33.0%) and 17.1% of women with fracture/osteoporosis had high antibody levels compared with 10.8% of women who had not ($P = 0.03$). Several types of pelvic/gynecologic surgery, including tubal sterilization, cervical conization, and cesarean section increased the likelihood of a positive antibody reaction and 47.2% of women who

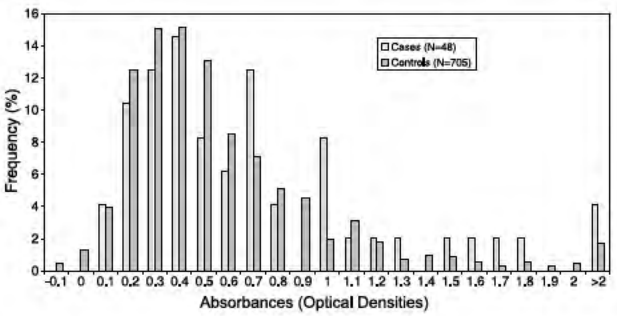

Figure 1. Distribution of absorbances from anti MUC1 antibody assay in cases with preoperative bloods and all controls.

Table 1. Occurrence of anti-MUC1 antibodies in control women by epidemiologic variables

	Negative	Positive*				Mean absorbance (SE)	Age adjusted *P*
		Low	Intermediate	High	Any		
All	467 (66.2)	94 (13.3)	57 (8.1)	87 (12.3)	238 (33.8)	0.44 (0.08)	
Age (y)							
<35	33 (50.0)	12 (18.2)	8 (12.1)	13 (19.7)	33 (50.0)	0.57 (0.24)	
35 44	90 (65.7)	17 (12.4)	18 (13.1)	12 (8.8)	47 (34.3)	0.45 (0.18)	
45 54	133 (67.5)	25 (12.7)	13 (6.6)	26 (13.2)	64 (32.5)	0.48 (0.13)	
55 64	118 (70.7)	24 (14.4)	11 (6.6)	14 (8.4)	49 (29.3)	0.37 (0.19)	
≥65	93 (67.4)	16 (11.6)	7 (5.1)†	22 (15.9)	45 (32.6)†	0.41 (0.23)	0.0009
Race							
White	454 (66.5)	92 (13.5)	55 (8.0)	82 (12.0)	229 (33.5)	0.44 (0.08)	
Non White	13 (59.1)	2 (9.1)	2 (9.1)	5 (22.7)	9 (40.9)	0.62 (0.37)	0.07
Religion							
Non Jewish	448 (66.3)	89 (13.2)	55 (8.1)	84 (12.4)	228 (33.7)	0.44 (0.09)	
Jewish	19 (65.5)	5 (17.2)	2 (6.9)	3 (10.3)	10 (34.5)	0.47 (0.35)	0.68
Marital status							
Never married	40 (66.7)	6 (10.0)	6 (10.0)	8 (13.3)	20 (33.3)	0.51 (0.22)	
Ever married	427 (66.2)	88 (13.6)	51 (7.9)	79 (12.2)	218 (33.8)	0.44 (0.09)	0.49
Pregnancy history							
Never pregnant	62 (66.7)	11 (11.8)	8 (8.6)	12 (12.9)	31 (33.3)	0.51 (0.19)	
Pregnant but no live births	20 (71.4)	4 (14.3)	3 (10.7)	1 (3.6)	8 (28.6)	0.45 (0.30)	
At least one live birth	385 (65.9)	79 (13.5)	46 (7.9)	74 (12.7)	199 (34.1)	0.43 (0.10)	0.24
Breast feeding (among parous women)							
Never breast fed	159 (66.0)	33 (13.7)	16 (6.6)	33 (13.7)	82 (34.0)	0.43 (0.15)	
Breast fed and no mastitis	211 (67.0)	41 (13.0)	29 (9.2)	34 (10.8)	104 (33.0)	0.42 (0.13)	
Breast fed and mastitis	15 (53.6)	5 (17.9)	1 (3.6)	7 (25.0)‡	13 (46.4)	0.55 (0.38)	0.95
OC use							
No	170 (70.8)	25 (10.4)	19 (7.9)	26 (10.8)	70 (29.2)	0.41 (0.15)	
Yes	297 (63.9)	69 (14.8)	38 (8.2)	61 (13.1)	168 (36.1)	0.46 (0.10)	0.42
OC use in premenopausal subjects							
No	39 (73.6)	4 (7.6)	8 (15.1)	2 (3.8)	14 (26.4)	0.45 (0.24)	
Yes	147 (59.3)	43 (17.3)‡	23 (9.3)	35 (14.1)†	101 (40.7)†	0.49 (0.13)	0.30
IUD use							
No	379 (66.4)	70 (12.3)	51 (8.9)	71 (12.4)	192 (33.6)	0.44 (0.09)	
Yes	88 (65.7)	24 (17.9)‡	6 (4.5)	16 (11.9)	46 (34.3)	0.45 (0.18)	0.60
Bone fracture/osteoporosis							
No	355 (67.0)	68 (12.8)	50 (9.4)	57 (10.8)	175 (33.0)	0.43 (0.10)	
Yes	112 (64.0)	26 (14.9)	7 (4.0)‡	30 (17.1)†	63 (36.0)	0.46 (0.16)	0.16
Colitis							
No	440 (65.9)	93 (13.9)	54 (8.1)	81 (12.1)	228 (34.1)	0.44 (0.08)	
Yes	27 (73.0)	1 (2.7)‡	3 (8.1)	6 (16.2)	10 (27.0)	0.44 (0.39)	0.97
Endometriosis							
No	430 (66.2)	86 (13.2)	53 (8.2)	81 (12.5)	220 (33.8)	0.45 (0.09)	
Yes	37 (67.3)	8 (14.6)	4 (7.3)	6 (10.9)	18 (32.7)	0.39 (0.32)	0.26
Pelvic surgery							
No pelvic surgery	304 (69.1)	52 (11.8)	33 (7.5)	51 (11.6)	136 (30.9)	0.43 (0.10)	
Hysterectomy only	26 (70.3)	4 (10.8)	1 (2.7)	6 (16.2)	11 (29.7)	0.39 (0.55)	
Tubal sterilization only	54 (64.3)	15 (17.9)	6 (7.1)	9 (10.7)	30 (35.7)	0.44 (0.23)	
Conization only	14 (58.3)	3 (12.5)	1 (4.2)	6 (25.0)	10 (41.7)	0.62 (0.34)	
Cesarean section only	31 (64.6)	5 (10.4)	6 (12.5)	6 (12.5)	17 (35.4)	0.47 (0.29)	
>1	38 (52.8)	15 (20.8)†	10 (13.9)†	9 (12.5)	34 (47.2)†	0.48 (0.27)	0.12
Smoking							
Never	228 (66.5)	43 (12.5)	30 (8.8)	42 (12.2)	115 (33.5)	0.45 (0.12)	
Former	182 (67.7)	36 (13.4)	20 (7.4)	31 (11.5)	87 (32.3)	0.43 (0.14)	
Current	57 (61.3)	15 (16.1)	7 (7.5)	14 (15.0)	36 (38.7)	0.45 (0.25)	0.86
Talc use							
None	208 (61.9)	43 (12.8)	41 (12.2)	44 (13.1)	128 (38.1)	0.46 (0.12)	
Body use only	117 (68.8)	23 (13.5)	9 (5.3)	21 (12.4)	53 (31.2)	0.46 (0.15)	
Genital use	142 (71.4)	28 (14.1)	7 (3.5)†	22 (11.1)	57 (28.6)†	0.39 (0.17)	0.08
No. conditions§							
0 or 1	119 (75.8)	15 (9.6)	10 (6.4)	13 (8.3)	38 (24.2)	0.37 (0.20)	
2	155 (67.1)	26 (11.3)	20 (8.7)	30 (13.0)	76 (32.9)	0.44 (0.14)	
3	114 (65.9)	25 (14.4)	17 (9.8)	17 (9.8)	59 (34.1)	0.45 (0.15)	
4	61 (57.0)	21 (19.6)	6 (5.6)	19 (17.8)	46 (43.0)	0.51 (0.20)	
5 or more	18 (48.6)	7 (18.9)†	4 (10.8)	8 (21.6)†	19 (51.4)‖	0.53 (0.41)	0.003

*Positive antibodies: low, $0.6 \geq A < 0.8$; intermediate, $0.8 \geq A < 1.0$; high $A \geq 1.0$; any $A \geq 0.6$.
†$P < 0.05$.
‡P between 0.05 and 0.15.
§Conditions include bone fracture/osteoporosis, mastitis, pelvic surgeries, IUD use, no genital talc use, OC use, and current smoking; Also adjusted for study center (Massachusetts, New Hampshire), parity (continuous), non-White race, and Jewish religion.
‖$P_{trend} = 0.0005$.

had more than one surgery had antibodies compared with 30.9% of women who never had pelvic surgery ($P = 0.01$). A surprising finding was that 38.1% of women who reported no use of cosmetic talc in hygiene had antibody compared with 28.6% of women who regularly used talc in genital hygiene ($P = 0.04$). The final entry shows the trend for elevated anti-MUC1 antibody levels by increasing number of antibody-promoting conditions. These included all variables significant in univariate analyses, such as OC use, bone fracture, mastitis, pelvic surgery, and genital talc use (where no use was considered the ''condition'') as well as variables of marginal significance in the univariate analysis, which nevertheless improved the overall model including current smoking and use of an intrauterine device (IUD). A significant trend ($P = 0.0005$) in the likelihood of having antibodies was observed such that 24.2% of women who had zero or one of condition had antibodies compared with 51.4% of women who had experienced five or more of these conditions.

Spearman (rank) correlations were calculated between the absorbance reading and several variables quantifiable on a numerical scale. No significant correlations were noted with number of live births, months of breast-feeding, or pack-years of smoking (data not shown). Weak but significant positive correlations were noted between absorbance values and months of OC use ($r = 0.09$, $P = 0.02$) and number of cesarean sections ($r = 0.10$, $P = 0.02$). A nonsignificant inverse correlation was noted between absorbance and estimated total applications of talc. When genital talc users were characterized by <weekly, weekly, and daily use, there was a trend of borderline significance ($P = 0.11$) for women using talc more frequently to have the lower antibody levels after adjustment for age, smoking, bone fractures, and OC or IUD use.

Risk for Ovarian Cancer Associated with Antibody-Promoting Events. The variables examined in relation to anti-MUC1 antibodies were then examined in relation to ovarian cancer risk, based upon case-control comparisons (Table 2). Odds ratios for ovarian cancer with each of these variables (except for age which was a matching variable) were calculated and adjusted for age, study site, exact parity, non-White race, and Jewish religion. Our study confirmed the influence of known ovarian cancer risk factors including parity, breast-feeding, and OC use. In addition, we identified previously unreported risk factors, including mastitis, relative risk (and 95% confidence limits) of 0.35 (0.16-0.77); IUD use, relative risk of 0.68 (0.50-0.91); and bone fracture, relative risk of 0.70 (0.53-0.91). The final entry shows the association between number of antibody-promoting conditions and ovarian cancer risk. Compared with women with zero or one condition, the risk for ovarian cancer decreased progressively with relative risks (and 95% confidence limits) of 0.69 (0.52-0.92), 0.64 (0.47-0.88), 0.49 (0.34-0.72), and 0.31 (0.16-0.61), respectively, for women with two, three, four, and five or more conditions ($P_{trend} < 0.0001$). This pattern mirrored the effect of these same conditions on the likelihood that control women had anti-MUC1 antibody (Fig. 2). Finally, risk by number of antibody-promoting conditions was examined separately for major histologic subtypes of ovarian cancer (Table 3). The inverse association was most evident for endometrioid cancers followed by undifferentiated and then invasive serous cancers. Numbers were too limited to make any definitive comments about predictors of antibodies among the 48 cases with preoperative bloods in whom anti-MUC1 antibodies were measured.

Discussion

To date, this is the largest study to examine determinants of anti-MUC1 antibodies and the first to show that conditions that generally increase the likelihood of having antibodies decrease the risk for ovarian cancer. MUC1 is normally present in a glycosylated, membrane-bound form on the apical surface of most polarized epithelial cells of the respiratory, genitourinary, and digestive tracts as well as breast ducts (12). With malignant transformation, epithelial cells lose polarity and overexpress MUC1 on their entire cell surface. A soluble, underglycosylated form circulates in cancer patients, thus becoming available for recognition by the immune system (6, 13). Some healthy women and men also have detectable MUC1 (albeit much lower levels) as well as anti-MUC1 antibodies. In women mostly ages 50 to 70 years, McGuckin et al. assessed the presence of circulating MUC1 using the cancer-associated serum antigen assay. Cancer-associated serum antigen concentrations were elevated in smokers and increased progressively with age (14). In a sample of women from the same study, Richards et al. then measured anti-MUC1 antibodies and found that virtually all women less than age 40 had antibodies and this percentage declined with age (4), somewhat similar to the pattern we observed. It is well established that women have MUC1 and anti-MUC1 antibodies during pregnancy and breast-feeding, presumably due to changes within the breast or uterus that alter MUC1 expression, glycosylation, or shedding (4, 15, 16). In addition, Hinoda et al. observed antibodies specific for the peptide backbone of MUC1 in patients with ulcerative colitis and speculated that inflammation may change MUC1 glycosylation and enhance its immunogenicity (17). One difficulty in evaluating these studies is that assays both for MUC1 and anti-MUC1 antibodies may differ. In measuring antibodies, assays will vary by the specific epitope of MUC1 and the secondary immunoglobulin antibody used. The assay in our study is based on the peptide backbone of MUC1 that we believe is closer to tumor MUC1 and we assessed total immunoglobulin levels including all isotypes, IgG, IgM, and IgA.

In our data, anti-MUC1 antibodies were associated with events affecting the reproductive tract, whose epithelia heavily express MUC1 (18). Injury and/or inflammation of these tissues, surgery, and other events might allow enhancement of MUC1 expression, spillage into circulation, and presentation to the immune system. Thus, the mechanism by which tubal sterilization reduces ovarian cancer risk, previously attributed to preventing substances like talc or endometrial cells from reaching the ovaries (19, 20), may include production of protective antibodies. In our data, cervical conization involving injury and repair of endocervical tissue was also associated with a nonsignificant increase in antibody formation and decrease in risk for ovarian cancer. Antibody formation was also directly correlated with number of cesarean sections, which involve incision and repair of the uterine wall and endometrium. Endometrial expression of MUC1 might also be affected by IUD use, as suggested by biopsies showing a low-grade, chronic inflammation with enhanced mucin staining (21). We found that IUD use increased the likelihood of antibodies in the ''low'' range and significantly decreased the risk for ovarian cancer. This is the first study to identify an inverse association between ovarian cancer and IUD use, whereas there is considerably more evidence that IUD use reduces risk for endometrial cancer (22), another tumor with high MUC1 expression (23).

An increased likelihood of MUC1-specific antibodies in the ''high'' range was found in women reporting bone fracture or a diagnosis of osteoporosis. Both conditions are known to be associated with high interleukin 6 levels (24, 25), an important regulatory cytokine for MUC1 expression (26). Furthermore, a bone fracture might be associated with release of hematopoetic precursors into the circulation, some of which may express MUC1 and be immunogenic (27). We also found an inverse association between bone fracture/osteoporosis and ovarian cancer risk, which to our knowledge has not been

shown previously. Interestingly, bone fracture is associated with reduced endometrial and breast cancer risk (28). Whereas this may simply reflect low estrogen, an influence of anti-MUC1 antibodies should also be considered. Besides bone fracture and IUD use, a third factor, which may link the etiology of ovarian and endometrial cancer, is smoking. A decreased risk for endometrial cancer is found in smokers, especially current smokers (29, 30). The data are less clear for ovarian cancer with two recent studies suggesting that smoking may increase the risk only for mucinous histologic subtypes (31, 32). Although current smoking was not clearly related to either anti-MUC1 antibody development or ovarian cancer risk in our univariate analyses, it did improve the cumulative index models in Tables 1 and 2. Furthermore, McGuckin's observation that smokers have higher serum MUC1 levels (presumably from damaged lung epithelium) provides a basis for linking current smoking to anti-MUC1 antibody production (14).

Table 2. Adjusted risk for ovarian cancer by epidemiologic variables in ovarian cancer cases and controls

	Cases	Controls	Adjusted odds ratio (95% confidence interval)	Adjusted P^*
Age (y)				
<35	68 (50.8)	66 (49.2)		
35 44	123 (46.2)	143 (53.8)		
45 54	198 (49.2)	204 (50.8)		
55 64	159 (48.6)	168 (51.4)		
≥65	120 (46.2)	140 (53.8)		
Race				
White	629 (47.4)	699 (52.6)	1.00	
Non White	39 (63.9)	22 (36.1)	2.14 (1.24 3.69)	0.007
Religion				
Non Jewish	614 (47.1)	690 (52.9)	1.00	
Jewish	54 (63.5)	31 (36.5)	1.88 (1.19 3.00)	0.007
Marital status				
Never married	107 (62.9)	63 (37.1)	1.00	
Ever married	561 (46.0)	658 (54.0)	0.66 (0.46 0.96)	0.03
Pregnancy history				
Never pregnant	169 (64.0)	95 (36.0)	1.00	
Pregnant but no live births	37 (54.4)	31 (45.6)	0.69 (0.40 1.18)	0.18
At least one live birth	462 (43.7)	595 (56.3)	0.41 (0.30 0.55)	<0.0001
Breast feeding (among parous women)				
Never breast fed	231 (48.6)	244 (51.4)	1.00	
Breast fed and no mastitis	219 (40.5)	322 (59.5)	0.72 (0.56 0.93)	0.01
Breast fed and mastitis	9 (23.7)	29 (76.3)	0.35 (0.16 0.77)	0.009
OC use				
No	312 (56.1)	244 (43.9)	1.00	
Yes	356 (42.7)	477 (57.3)	0.55 (0.43 0.69)	<0.0001
OC use in premenopausal subjects				
No	102 (65.4)	54 (34.6)	1.00	
Yes	184 (41.6)	258 (58.4)	0.38 (0.26 0.57)	<0.0001
IUD use				
No	577 (49.6)	586 (50.4)	1.00	
Yes	91 (40.3)	135 (59.7)	0.68 (0.50 0.91)	0.01
Bone fracture/osteoporosis				
No	544 (50.0)	545 (50.0)	1.00	
Yes	124 (41.3)	176 (58.7)	0.70 (0.53 0.91)	0.007
Colitis				
No	645 (48.6)	683 (51.4)	1.00	
Yes	23 (37.7)	38 (62.3)	0.58 (0.34 1.00)	0.05
Endometriosis				
No	611 (47.8)	666 (52.2)	1.00	
Yes	57 (50.9)	55 (49.1)	1.02 (0.69 1.52)	0.92
Pelvic surgery				
No pelvic surgery	441 (49.4)	452 (50.6)	1.00	
Hysterectomy only	39 (51.3)	37 (48.7)	1.23 (0.75 2.03)	0.40
Tubal sterilization only	69 (44.5)	86 (55.5)	1.00 (0.70 1.42)	0.98
Conization only	15 (38.5)	24 (61.5)	0.64 (0.33 1.26)	0.20
Cesarean section only	60 (55.0)	49 (45.0)	1.38 (0.91 2.09)	0.13
>1 surgery	44 (37.6)	73 (62.4)	0.73 (0.48 1.10)	0.13
Smoking				
Never	338 (48.9)	353 (51.1)	1.00	
Former	235 (46.2)	274 (53.8)	1.15 (0.83 1.60)	0.40
Current	95 (50.3)	94 (49.7)	0.94 (0.74 1.19)	0.59
Talc use				
None	319 (48.2)	343 (51.8)	1.00	
Body use only	135 (43.7)	174 (56.3)	0.87 (0.66 1.15)	0.33
Genital use	214 (51.2)	204 (48.8)	1.16 (0.90 1.49)	0.25
No. conditions[†]				
0 or 1	218 (57.4)	162 (42.6)	1.00	
2	220 (48.1)	237 (51.9)	0.69 (0.52 0.92)	0.01
3	150 (45.9)	177 (54.1)	0.64 (0.47 0.88)	0.005
4	67 (38.3)	108 (61.7)	0.49 (0.34 0.72)	0.0002
5 or more	13 (26.0)	37 (74.0)	0.31 (0.16 0.61)	0.0007

*Adjusted for age (continuous), study center (Massachusetts, New Hampshire), parity (continuous), non-White race, and Jewish religion.
[†]Conditions include bone fracture/osteoporosis, mastitis, pelvic surgeries, IUD use, no genital talc use, OC use, and current smoking.

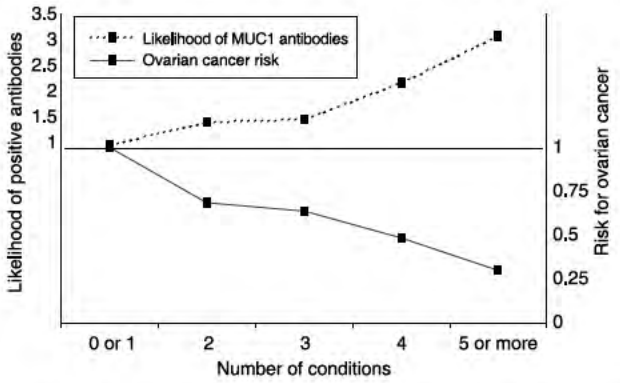

Figure 2. Likelihood of anti MUC1 antibodies by index of number of conditions and risk for ovarian cancer by same index.

OC use is a strong protective factor for ovarian (and endometrial) cancer and also seemed to generate anti-MUC1 antibodies, particularly among premenopausal women. CA15-3 (MUC1) levels in saliva were found to be 75% higher in OC users compared with nonusers, a nonsignificant difference in that small study (33). Other studies suggest that MUC1 expression in the endometrium is progesterone dependent (34) and up-regulated by exogenous progesterone (35). Considered together, these observations support the speculation that OC users may have higher MUC1 levels that could translate into higher antibody production.

History of mastitis was associated with both increased anti-MUC1 antibodies and decreased ovarian cancer risk in our study. We believe this is an important finding in light of our previous report of a long-term breast cancer survivor in whom MUC1-specific antibody production and mucin-specific T lymphocytes became elevated following mastitis in pregnancy (36). The lactating breast secretes a form of MUC1 that is similar to the underglycosylated form of MUC1 produced by tumors. Thus, mastitis may lead to a potent anti-MUC1 and antitumor immune response, which could explain the substantial decreased risk for ovarian cancer associated with mastitis found in our current study.

Curiously, we found that use of talc in the genital area was associated with significantly decreased levels of anti-MUC1 antibodies. Use of talc in the genital area would expose at least lower genital tract epithelia to talc and conceivably affect MUC1 expression in these tissues. In serial assays of pleural fluid in patients who received talc pleurodesis, inflammatory mediators eventually became depressed (37). Use of talc in the genital area has been consistently found to increase the risk for ovarian cancer in several meta-analyses (38-40). However, some investigators have challenged the association because of uncertainty about its biological basis and the absence of a dose-response relationship (38, 40). Although our present finding may also meet with skepticism, a testable hypothesis is now suggested by the possible link between genital talc exposure and systemic diminution of anti-MUC1 antibodies.

Existing theories of ovarian cancer pathogenesis have invoked incessant ovulation, gonadotropin excess, androgen excess, progesterone deficiency, or deleterious effects of inflammation to explain risk factors for ovarian cancer (41-44). Our findings offer an additional perspective on how OC use, tubal sterilization, and even talc use might exert their effects on ovarian cancer risk and suggests an entirely new set of protective factors such as mastitis, IUD use, and bone fracture that might be explained by the same immune-mediated mechanism. Interestingly, this mechanism may also explain the decreased risk for ovarian cancer associated with mumps parotitis noted in older studies conducted before the widespread use of vaccination (45, 46). Analogous to mastitis, infection of MUC1-rich salivary glands might also lead to an anti-MUC1 immune response and antibody production. Clearly, we have not explained all features of ovarian cancer including the "dose-related" decrease in risk associated with multiple pregnancies and length of breast-feeding. Based on the studies reporting anti-MUC1 antibodies in women currently pregnant or breast-feeding, we had expected, but did not observe, that antibodies would increase with the more pregnancies a woman had or the longer she breast-fed. However, it should also be appreciated that anti-MUC1 antibodies are just one of several immune-effecter mechanisms that may also include helper and cytotoxic MUC1-specific T cells that are generated by MUC1 presentation to the immune system. Indeed, the reactions described in sera and T cells from multiparous women suggest that a complete picture of the link between ovarian cancer risk and MUC1 immunity will require assessment of cell-mediated reactions. In addition, immunity to other human mucins, including MUC16 (CA 125), may also need to be examined.

The principal limitation of our study comes from its case-control design. Exposure information was collected by self-report after the diagnosis in cases, introducing the possibility of misclassification. More importantly, we were unable to directly compare anti-MUC1 antibody levels in cases and controls and directly calculate odds ratios based on antibody levels because the cancer itself leads to production of antibodies. Consequently, assessing antibodies in cases after the diagnosis is not useful for identifying earlier events that influenced antibody generation or the predictive value of such antibodies. Prospective studies, in which blood samples are obtained decades or years before the development of ovarian cancer, will be necessary to assess directly the predictive value of anti-MUC1 antibodies on ovarian cancer risk. In addition, prospective studies before and after events like tubal sterilization, IUD use, mastitis, etc. that document the precise changes in the status of anti-MUC1 antibodies will refine our "cumulative index model" with its crude assumption that all events might be of equal potency in ability to generate

Table 3. Adjusted risks, 95% confidence intervals, and trends for ovarian cancer of different histologic types associated with number of conditions predisposing to MUC1 antibodies

No. conditions	Histologic subtype				
	Serous borderline (n = 91)	Serous invasive (n = 261)	Mucinous (n = 73)	Endometrioid/CC (n = 201)	Other/Undiff. (n = 42)
0 or 1	1.00	1.00	1.00	1.00	1.00
2	0.74 (0.40 1.35)	0.80 (0.55 1.17)	0.62 (0.32 1.18)	0.68 (0.45 1.02)	0.57 (0.27 1.22)
3	0.72 (0.38 1.38)	0.91 (0.61 1.37)	0.41 (0.19 0.89)	0.62 (0.39 0.98)	0.32 (0.12 0.84)
4	0.85 (0.40 1.78)	0.53 (0.31 0.90)	0.80 (0.36 1.75)	0.38 (0.20 0.71)	0.27 (0.08 0.97)
5 or more	0.38 (0.08 1.74)	0.38 (0.14 1.02)	0.65 (0.18 2.37)	0.17 (0.04 0.73)	0.31 (0.04 2.42)
P_{trend}	0.32	0.02	0.30	0.0002	0.008

NOTE: Adjusted for age (continuous), study site (Massachusetts, New Hampshire), ethnicity (white, non-White), religious background (Jewish, non-Jewish), and parity (continuous).

antibodies. Thus, we make no claim this model is final but rather represents a simple foundation for a paradigm shift that will incorporate MUC1 immunity as a key mechanism through which many risk factors for ovarian cancer may exert their influence.

In summary, evidence from this case-control study of ovarian cancer suggests that events predicting the presence of anti-MUC1 antibodies are inversely associated with ovarian cancer risk and that the more conditions a woman experienced to raise antibodies the lower is her risk for ovarian cancer. We believe these data support the immune response as one mechanism of action of ''traditional'' ovarian cancer risk factors such as OC use and tubal sterilization, as well as novel ones observed in this study including mastitis, bone fracture, and IUD use. If, as we would like to propose, the immune response is a major mechanism, the implications are profound because it may eventually offer new avenues for ovarian cancer prevention through vaccines to stimulate immunity against MUC1 and perhaps other antigens expressed in ovarian cancer. Much work would need to be done, including prospective documentation of the precise changes in cell-mediated and humoral responses to MUC1 associated with pregnancy, breast-feeding, mastitis, and other events. Such studies may have implications beyond ovarian cancer and apply to other cancers with high MUC1 expression including endometrial and breast cancer.

Acknowledgments

Special thanks to Mr. Abbott Dean in memory of his wife, Sandra, and to all the participants of this study.

References

1. Ho SB, Niehans GA, Lyftogt C, et al. Heterogeneity of mucin gene expression in normal and neoplastic tissues. Cancer Res 1993;53:641 51.
2. Kotera Y, Fontenot JD, Pecher G, Metzgar RS, Finn OJ. Humoral immunity against a tandem repeat epitope of human mucin MUC-1 in sera from breast, pancreatic, and colon cancer patients. Cancer Res 1994;54:2856 60.
3. von Mensdorff-Pouilly S, Gourevitch MM, Kenemans P, et al. Humoral immune response to polymorphic epithelial mucin (MUC-1) in patients with benign and malignant breast tumours. Eur J Cancer 1996;32:1325 31.
4. Richards ER, Devine PL, Quin RJ, Fontenot JD, Ward BG, McGuckin MA. Antibodies reactive with the protein core of MUC1 mucin are present in ovarian cancer patients and healthy women. Cancer Immunol Immunother 1998;46:245 52.
5. Hamanaka Y, Suehiro Y, Fukui M, Shikichi K, Imai K, Hinoda Y. Circulating anti-MUC1 IgG antibodies as a favorable prognostic factor for pancreatic cancer. Int J Cancer 2003;103:97 100.
6. Vlad AD, Kettel JC, Alajez NM, Carlos CA, Finn OJ. MUC1 immunobiology: from discovery to clinical applications. Adv Immunol 2004;82:249 93.
7. Agrawal B, Reddish MA, Krantz MJ, Longenecker BM. Does pregnancy immunize against breast cancer? Cancer Res 1995;55:2257 61.
8. Janerich DT. The influence of pregnancy on breast cancer risk: is it endocrinological or immunological? Med Hypotheses 1980;6:1149 55.
9. Forsman LM, Jouppila PI, Andersson LC. Sera from multiparous women contain antibodies mediating cytotoxicity against breast carcinoma cells. Scand J Immunol 1984;19:135 9.
10. Shields LB, Gercel-Taylor C, Yashar CM, et al. Induction of immune responses to ovarian tumor antigens by multiparity. J Soc Gynecol Investig 1997;4:298 304.
11. Agrawal B, Reddish MA, Longenecker BM. *In vitro* induction of MUC-1 peptide-specific type 1 T lymphocyte and cytotoxic T lymphocyte responses from healthy multiparous donors. J Immunol 1996;157:2089 95.
12. Gendler SJ, Spicer AP. Epithelial mucin genes. Annu Rev Physiol 1995;57:607 34.
13. Fontenot JD, Mariappan SV, Catasti P, Domenech N, Finn OJ, Gupta G. Structure of a tumor associated antigen containing a tandemly repeated immunodominance epitope. J Biomol Struct Dyn 1995;3:245 60.
14. McGuckin MA, Ramm LE, Joy GJ, Devine PL, Ward BG. Circulating tumor associated mucin concentrations determined by the CASA assay in healthy women. Clin Chim Acta 1993;214:139 51.
15. Bon GG, Kenemans P, Verstraeten AA, et al. Maternal serum Ca125 and Ca 15-3 antigen levels in normal and pathological pregnancy. Fetal Diagn Ther 2001;16:166 72.
16. Croce MV, Isla-Larrain MT, Price MR, Segal-Eiras A. Detection of circulating mammary mucin (Muc1) and MUC1 immune complexes (Muc1-CIC) in healthy women. Int J Biol Markers 2001;16:112 20.
17. Hinoda Y, Nakagawa N, Nakamura H, et al. Detection of a circulating antibody against a peptide epitope on a mucin core protein, MUC1, in ulcerative colitis. Immunol Lett 1993;35:163 8.
18. Gipson IK, Ho SB, Spurr-Michaud SJ, et al. Mucin genes expressed by human female reproductive tract epithelia. Biol Reprod 1997;56:999 1011.
19. Hankinson SE, Hunter DJ, Colditz GA. Tubal ligation, hysterectomy, and risk of ovarian cancer. JAMA 1993;270:2813 8.
20. Green A, Purdie PD, Bain C, et al. Tubal sterilization, hysterectomy and decreased risk of ovarian cancer. Survey of Women's Health Study Group. Int J Cancer 1997;71:948 51.
21. Hester LL Jr, Kellett WW III, Spicer SS, Williamson HO, Pratt-Thomas HR. Effects of the intrauterine contraceptive device on endometrial enzyme and carbohydrate histochemistry. Am J Obstet Gynecol 1970;106:1144 54.
22. Hubacher D, Grimes DA. Noncontraceptive health benefits of intrauterine devices: a systematic review. Obstet Gynecol Surv 2002;57:120 8.
23. Sivridis E, Giatromanolaki A, Koukourakis MI, Georgiou L, Anastasiadis P. Patterns of episialin/MUC1 expression in endometrial carcinomas and prognostic relevance. Histopathology 2002;40:92 100.
24. Papadopoulos NG, Georganas K, Skoutellas V, Konstantellos E, Lyritis GP. Correlation of interueukin-6 serum levels with bone density in postmenopausal women. Clin Rheumatol 1997;16:162 5.
25. Strecker W, Gebhard F, Rager J, Bruckner UB, Steinbach G, Kinzl L. Early biochemical characterization of soft-tissue trauma and fracture trauma. J Trauma Injury Infect Crit Care 1999;47:358 64.
26. Gaemers IC, Vox HL, Volders HH, van der Valk SW, Hilkens J. A stat-responsive element in the promoter of the episialin/MUC1 gene is involved in its overexpression in carcinoma cells. J Biol Chem 2001;276:6191 9.
27. Brugger W, Buhring HJ, Grunebach F, et al. Expression of MUC-1 epitopes on normal bone marrow: implications for the detection of micrometastatic tumor cells. J Clin Oncol 1999;17:1535 44.
28. Newcomb PA, Trentham-Dietz A, Egan KM, et al. Fracture history and risk of breast and endometrial cancer. Am J Epidemiol 2001;153:1071 8.
29. Terry PD, Miller AB, Rohan TE. A prospective cohort study of cigarette smoking and the risk of endometrial cancer. Br J Cancer 2002;86:1430 5.
30. Brinton LA, Barrett RJ, Berman ML, Mortel R, Twiggs LB, Wilbanks GD. Cigarette smoking and the risk of endometrial cancer. Am J Epidemiol 1993;137:281 91.
31. Pan SY, Ugnat AM, Mao Y, Wen SW, Johnson KC. Association of cigarette smoking and the risk of ovarian cancer. Int J Cancer 2004;111:124 30.
32. Zhang Y, Coogan PF, Palmer JR, Strom BL, Rosenberg L. Cigarette smoking and increased risk of mucinous epithelial ovarian cancer. Am J Epidemiol 2004;159:33 9.
33. McIntyre R, Bigler L, Dellinger T, Pfeifer M, Mannery T, Stredkfus C. Oral contraceptive usage and the expression of CA 15-3 and c-erB-2 in the saliva of healthy women. Oral Surg Med Pathol Radiol Endod 1999;88:687 90.
34. Hild-Petito S, Fazleabas AT, Julian J, Carson DD. Mucin (Muc-1) expression is differentially regulated in uterine luminal and glandular epithelia of the baboon (*Papio anubis*). Biol Reprod 1993;54:939 47.
35. Hewetson A, Chilton BS. Molecular cloning and hormone-dependent expression of rabbit Muc1 and the cervix and uterus. Biol Reprod 1997;57:468 77.
36. Jerome KR, Kir AD, Pecher G, Ferguson WW, Finn OJ. A survivor of breast cancer with immunity to MUC-1 mucin, and lactational mastitis. Cancer Immunol Immunother 1997;43:355 60.
37. D'Agostino P, Camemi AR, Arcoleo F, et al. Matrix metalloproteinases production in malignant pleural effusions after talc pleuodesis. Clin Exp Immunol 2003;34:138 42.
38. Gross AJ, Berg PH. A meta-analytical approach examining the potential relationship between talc exposure and ovarian cancer. J Expo Anal Environ Epidemiol 1995;5:181 95.
39. Cramer DW, Liberman RF, Titus-Ernstoff L, et al. Genital talc exposure and risk of ovarian cancer. Int J Cancer 1999;81:351 6.
40. Huncharek M, Gerschwind JF, Kupelnick B. Perineal application of cosmetic talc and risk of invasive epithelial ovarian cancer: a meta-analysis of 11,933 subjects from sixteen observational studies. Anticancer Res 2003;23:1955 60.
41. Fathalla MF. Incessant ovulation: a factor in ovarian neoplasia? Lancet 1971;2:163.
42. Cramer DE, Welch WR. Determinants of ovarian cancer risk. II. Inference regarding pathogenesis. Natl Cancer Inst 1983;1:717 21.
43. Ness RB, Cottreau C. Possible role of ovarian epithelial inflammation in ovarian cancer. J Natl Cancer Inst 1999;91:1459 67.
44. Risch HA. Hormonal etiology of epithelial ovarian cancer, with a hypothesis concerning the role of androgens and progesterone. J Natl Cancer Inst 1998;90:1774 86.
45. West RD. Epidemiologic study of malignancies of the ovaries. Cancer 1966;19:1001.
46. Newhouse ML, Pearson RM, Fullerton JM, Boesen EA, Shannon HS. A case control study of carcinoma of the ovary. Br J Prev Soc Med 1977;31:148 53.

Exhibit 112

Published OnlineFirst April 20, 2010; DOI: 10.1158/1055-9965.EPI-09-1221

Research Article

Cancer Epidemiology, Biomarkers & Prevention

Perineal Use of Talcum Powder and Endometrial Cancer Risk

Stalo Karageorgi[1,2], Margaret A. Gates[3,4], Susan E. Hankinson[3,4], and Immaculata De Vivo[2,3,4]

Abstract

Background: Several studies have reported a positive association between perineal use of talcum powder among adult women and ovarian cancer risk. However, the relationship between talcum powder use and other gynecologic malignancies such as endometrial cancer has not been examined, and little information is available on nonhormonal risk factors for endometrial cancer.

Methods: Perineal use of talcum powder was assessed in 1982 in the Nurses' Health Study. Approximately 40% of women who responded to the questions about perineal use of talcum powder reported ever use. Cox proportional hazards models were used to estimate the incidence rate ratio of endometrial cancer and 95% confidence interval (CI), adjusted for body mass index and other potential confounders. We evaluated the relationship among all women and stratified by menopausal status.

Results: Our analysis included 66,028 women with 599 incident cases of invasive endometrial adenocarcinoma diagnosed between 1982 and 2004. Although no association was observed overall, the association varied by menopausal status (P interaction = 0.02) and a positive association was observed among postmenopausal women; ever use of talcum powder was associated with a 21% increase in risk of endometrial cancer (95% CI, 1.02 1.44), whereas regular use (at least once a week) was associated with a 24% increase in risk (95% CI, 1.03 1.48). In addition, we observed a borderline increase in risk with increasing frequency of use (P trend = 0.04).

Conclusions: Our results suggest that perineal talcum powder use increases the risk of endometrial cancer, particularly among postmenopausal women.

Impact: Future and larger studies are needed to confirm this association and investigate potential mechanisms. *Cancer Epidemiol Biomarkers Prev; 19(5); 1269 75. ©2010 AACR.*

Introduction

Several studies have reported a positive association between use of talcum powder on the perineal area and ovarian cancer risk (1, 2). In 2006, the IARC classified perineal use of talc as a possible carcinogen (2). In a meta-analysis, data from 16 studies suggested that talc may increase ovarian cancer risk by 30% (1). However, no previous studies have investigated whether talcum powder applied to the perineal area was associated with other gynecologic malignancies such as endometrial cancer. Furthermore, little information is available on factors that influence risk of endometrial cancer, a hormone-responsive cancer (3), through nonhormonal pathways.

Authors' Affiliations: [1]Department of Environmental Health, [2]Program in Molecular and Genetic Epidemiology, [3]Department of Epidemiology, Harvard School of Public Health, and [4]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

Corresponding Author: Stalo Karageorgi, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115. Phone: 617 935 3617; Fax: 617 432 1722. E mail: skarageo@hsph.harvard.edu

doi: 10.1158/1055 9965.EPI 09 1221

Talc is a hydrous magnesium silicate mineral chemically similar to the serpentine class of asbestos (4). In nature, talc is commonly found with other minerals such as chlorite, carbonates, amphiboles, and serpentines (5), in a fibrous or nonfibrous foliated structure (6). Before 1976, talcum powder was commonly contaminated with asbestos due to the proximity of talc and asbestos deposits in nature (7). Guidelines were set thereafter in the United States to ensure that only talc with no detectable levels of asbestos was used in cosmetic products (8). Talc used in powders is finely ground (4, 9); however, it is unknown whether processing of talc makes it more hazardous or increases its potential carcinogenicity (10, 11).

Older studies suggested a link between asbestos exposure in female workers and ovarian cancer incidence (8), which, together with the pathologic similarity between malignant pleural mesothelioma and ovarian tumors and the evidence of talc particles found in ovarian tissue (12), led to the investigation of whether talcum powder increased the risk of ovarian cancer. Although perineal talc use is common among adults as many as 40% of women in the United States have used talcum powder for feminine hygiene (13, 14) additional studies are needed to assess other possible health consequences. In this analysis, we used data from the Nurses' Health

Published OnlineFirst April 20, 2010; DOI: 10.1158/1055-9965.EPI-09-1221

Study (NHS) to assess whether genital use of talcum powder among women confers an increased risk of endometrial cancer.

Materials and Methods

Study population. The NHS is a prospective cohort study established in 1976, when 121,700 married female registered nurses residing in 11 U.S. states and were between the ages of 30 to 55 completed a baseline mailed questionnaire inquiring about various disease exposures and personal health status. Revised questionnaires were mailed biennially to update exposure and disease information. The follow-up rate through 2004, as measured as a percentage of total possible person-years, was 95.5%. Deaths in this cohort were identified by next-of-kin reports, the U.S. Postal Service, or through searches of the National Death Index (15). The Committee on the Use of Human Subjects in Research at Brigham and Women's Hospital, Boston, MA approved this analysis.

For this study, we excluded women who had had a hysterectomy (n = 30,287), had surgical menopause (n = 123), reported endometrial cancer (n = 89), reported any other type of cancer excluding nonmelanoma skin cancer (n = 1,422), or had died (n = 1,203) prior to assessment of talcum powder use in 1982. Women who were missing body mass index (BMI) for at least two consecutive cycles prior to and including the 1982 cycle were temporarily excluded until they reported their weight again. A total of 66,028 women remained for analysis.

Diagnosis of endometrial cancer cases. Information on endometrial cancer diagnoses was collected beginning in 1978 and at each subsequent questionnaire cycle. Women who reported a diagnosis of endometrial cancer were asked for permission to review their medical records. Only cases of invasive type I endometrioid adenocarcinoma (International Classification of Diseases for Oncology-3 histology codes 8380-83) confirmed by medical records were included in this analysis.

Data collection. Use of talcum powder was assessed in 1982. Participants were asked whether they had ever commonly applied talcum, baby, or deodorizing powder to the perineal area (no, less than once a week, one to six times a week, daily), or to sanitary napkins (no, yes).

Age at menarche was collected at baseline in 1976. Menopausal status, age at menopause and type of menopause were collected at baseline and every 2 years thereafter. On each questionnaire, women were asked whether their menstrual periods had ceased permanently [yes, no longer have periods; yes, periods induced by hormones (asked from 1988 onwards); no, unsure]. Women who responded "yes" were classified as postmenopausal and this status was carried forward into all future cycles. Women who responded "no or unsure" were classified as premenopausal. Women missing menopausal status were classified as postmenopausal if they were above a certain age (>54 years for current smokers, >56 years for former or nonsmokers). Nurses were also asked at what age their periods ceased and for what reason (natural, surgery). Age at menopause in the NHS is reported with a high degree of reproducibility and accuracy (16).

Women reported the number of pregnancies lasting 6 months or more and their age at first birth on every questionnaire through 1984. In 1996, they were asked to report their lifetime pregnancy history. These data were used to derive biennially updated variables for parity, age at first birth, and age at last birth.

Postmenopausal hormone (PMH) use and duration of use were first asked in 1976. Beginning in 1978 and at each 2-year follow-up, women were asked whether they currently used PMH, the number of months used during the 24 months prior to the questionnaire, and the type of PMH. Duration of ever hormone use was calculated as the cumulative duration of all types of PMH use reported over the follow-up.

Information on weight, diabetes, and smoking were collected at baseline and updated every 2 years. BMI (kg/m^2) was calculated using height reported in 1976 and weight reported at each cycle. Height and weight are accurately reported in the NHS (17). Because BMI is an important confounder, we carried forward BMI from the prior cycle for women missing weight in one cycle. If BMI was missing for two consecutive cycles, we excluded the person-time for these women until they reported their weight again. Nurses were asked whether they were past or current cigarette smokers and the number of cigarettes smoked per day. Pack-years were calculated by multiplying smoking duration in years by packs of cigarettes smoked per day.

First-degree family history of endometrial cancer was collected only in 1996. Information on oral contraceptive (OC) use and duration of use in months was collected every 2 years until 1982, at which time fewer than 500 women reported using OCs (18).

Statistical analysis. We used multivariate Cox proportional hazards models stratified by age in months at the start of follow-up and calendar year to estimate incidence rate ratios (RR) and 95% confidence intervals (CI). The time scale used was follow-up time in months, which is equivalent to using age in months as the time scale. Participants were followed from the age in months at the date of return of the 1982 questionnaire until the end of the study (June 1, 2004). Women contributed person-time until age in months at death, diagnosis of endometrial cancer, report of any other cancer excluding nonmelanoma skin cancer, report of hysterectomy, report of surgical menopause (one or two ovaries removed), loss to follow-up, or the end of the study, whichever came first. We evaluated the associations among all women and stratified by menopausal status. In the analysis among postmenopausal women, women started contributing person-times after they became menopausal.

Table 1. Age and age-standardized baseline characteristics according to perineal talc use in 1982 among 66,088 women in the NHS

Characteristics	Ever perineal talc use	
	No (n = 40,958)	Yes (n = 25,130)
Age	48	48
Age at menarche	12.6	12.5
Nulliparous (%)	6.0	5.6
Parity (mean)*	3.2	3.2
Age at first birth*	25	25
Age at last birth*	31	31
Ever OC use (%)	47	46
OC duration (mo)†	52	51
Postmenopausal (%)	40	40
Age at menopause‡	49	49
Ever PMH use (%)‡	30	29
PMH duration (mo)†,‡	35	34
Ever cigarette smoking (%)	57	56
Pack-years†	22	21
BMI (kg/m^2)	24.2	25.6
BMI (kg/m^2) [%]		
<25	67.9	55.8
25-29	22.6	27.2
≥30	9.5	17.0
Diabetes (%)	0.8	0.9
Family history uterine cancer (%)	3.0	2.9
IUD (%)	2.9	2.8
Diaphragm (%)	4.2	4.4

NOTE: Characteristics adjusted for age in 5-y categories (<45, 45-49, 50-54, 55-59, 60-64, >65).
*Among parous women only.
†Among users.
‡Among postmenopausal women.

Talcum powder use was modeled as ever use (no, yes), regular use (at least once a week), frequency of use (0, less than once a week, one to six times a week, daily use), and indirect use on sanitary napkins (no, yes). To test for trend, we weighted the categories of frequency of perineal talc use as 0, 2, 15.5, and 30 days of use per month and calculated the Wald test. The final model was adjusted for age at menarche, age at menopause, parity, age at last birth, PMH use duration, OC use duration, BMI, smoking pack-years, report of diabetes, and family history of endometrial cancer. All variables in the model were entered as time-dependent and updated biennially at the date in months of return of each questionnaire, with the exception of talcum powder use, age at menarche, and family history of endometrial cancer, which were collected at a single time point (in 1982, 1976, and 1996, respectively) and were entered as baseline variables. Other than BMI, which was modeled continuously, all exposures and covariates were categorized and an indicator variable was created for each category (see tables for categories). For covariates with missing data, a missing indicator was included in the model.

Results

Our analysis included 66,028 women with 599 incident cases of confirmed endometrial cancer diagnosed between 1982 and 2004. A total of 1,069,130 person-years were accumulated over 22 years of follow-up. The mean age at the start of follow-up was 48 years, and women were followed for an average of 16 years. Ever users of perineal talc and never users were similar in terms of their baseline characteristics, except for BMI (Table 1). Women who reported ever using talcum powder were more likely to be obese than never users (17% versus 10%), and talc users had a higher mean BMI (25.6 versus 24.2 kg/m^2). In addition, users were less likely to be nulliparous (5.6% versus 6.0%).

After control for confounding, ever use of perineal talcum powder was associated with a borderline significant 13% increase in endometrial cancer risk among all

Published OnlineFirst April 20, 2010; DOI: 10.1158/1055-9965.EPI-09-1221

Table 2. Incidence RRs and 95% CIs for ever and regular talc use and endometrial cancer risk among all women and stratified by menopausal status in the NHS

Women	Ever perineal talc use		Regular perineal talc use (at least once a week)	
	No	Yes	No	Yes
All women				
Cases	334	265	397	202
Person-years	687,327	420,106	806,391	301,041
RR (95% CI)*	1.00	1.13 (0.96-1.33)	1.00	1.17 (0.99-1.40)
Postmenopausal				
Cases	287	242	344	185
Person-years	461,381	281,958	538,227	205,113
RR (95% CI)*	1.00	1.21 (1.02-1.44)	1.00	1.24 (1.03-1.48)
Premenopausal				
Cases	47	23	53	17
Person-years	204,180	125,414	242,419	87,176
RR (95% CI)†	1.00	0.69 (0.40-1.19)	1.00	0.77 (0.42-1.39)
P interaction‡		0.02		0.07

*Adjusted for age, parity (0, 1, 2, 3, 4+), age at last birth (nulliparous, <30, 30-34, 35-39, ≥40), age at menarche (≤11, 12, 13, ≥14), age at menopause (premenopausal, <45, 45-49, 50-54, ≥55), OC duration (never, ≤12 mo, 13-36 mo, 37-72 mo, >72 mo), PMH duration (premenopausal/never, past <5 y, past 5+ years, current <5 y, current 5+ y), BMI (continuous), smoking pack-years (0, 1-20, 21-40, 40+), diabetes (no, yes), and family history of uterine cancer (no, yes); also adjusted for menopausal status (premenopausal, postmenopausal) among all women only.

†Adjusted for age, parity (0, 1, 2, 3, 4+), age at last birth (nulliparous, <30, 30-34, 35-39, ≥40), age at menarche (≤11, 12, 13, ≥14), OC duration (never, ≤12 mo, 13-36 mo, 37-72 mo, >72 mo), BMI (continuous), smoking pack-years (0, 1-20, 21-40, 40+), diabetes (no, yes), and family history of uterine cancer (no, yes).

‡Wald test for interaction term between talc use and menopausal status.

women, and a statistically significant 21% increase in risk among postmenopausal women (95% CI, 1.02-1.44; Table 2). There was no evidence of an association among premenopausal women, although the confidence interval was wide due to the small number of premenopausal cases. Because of significant differences in the results by menopausal status (*P* interaction = 0.02), as well as the small number of premenopausal cases, further detailed analyses of talc and risk were conducted among postmenopausal women only (Table 3). When we examined the association between frequency of perineal talc use and risk among postmenopausal women, there was a borderline trend of increasing risk with increasing frequency of use (*P* trend = 0.04; Table 3); in addition, the risks associated with perineal talc use one to six times a week or daily were elevated and borderline statistically significant. Regular use of talcum powder, defined as use at least once a week, was associated with a 24% increase in risk among postmenopausal women (95% CI, 1.03-1.48). The difference between the age-adjusted and multivariate results was due to confounding by BMI. Indirect use on sanitary napkins was not associated with risk. We further restricted the analyses to a group of postmenopausal women at low risk for endometrial cancer (19), consisting of normal-weight women with no history of PMH use (*n* = 27 cases). The association with regular use of talcum powder was maintained in this low-risk group despite decreased power (BMI < 25 never PMH users, cases = 27: RR, 2.44; 95% CI, 1.02-5.80; BMI ≥ 25 never PMH users, cases = 157: RR, 1.52; 95% CI, 1.09-2.12; BMI < 25 ever PMH users, cases = 139: RR, 0.90; 95% CI, 0.59-1.36; BMI ≥ 25 ever PMH users, cases = 170: RR, 1.27; 95% CI, 0.92-1.75; data not shown).

Discussion

We report for the first time an association between perineal use of talcum powder and endometrial cancer risk. In this large prospective study, we found a significant although modest increase in risk for endometrial cancer among postmenopausal women with a history of perineal use of talcum powder. The presence of an association among postmenopausal but not premenopausal women may be attributed to a longer duration of exposure or an increased latency in postmenopausal women, or to pathologic differences between premenopausal and postmenopausal endometrial cancer. The strength of the association is comparable with that reported between talc and ovarian cancer risk, a relationship investigated since the early 1980s (1, 20). In addition, some studies have

Published OnlineFirst April 20, 2010; DOI: 10.1158/1055-9965.EPI-09-1221

reported a positive association between talc and the serous (21) or endometrioid histologic subtype of ovarian cancer (22, 23), the latter of which resembles endometrial carcinoma, providing additional support for our findings.

We observed a borderline increase in risk with increasing frequency of talc use, indicating a possible dose-response. In addition, the association with regular talc use was slightly stronger than the association for ever use. If regular users recall use more accurately than non-regular users, regular use of talcum powder may be a more precise classification of exposure. Alternatively, regular use may be necessary to cause an adverse effect, if there is a threshold level of exposure below which carcinogenicity is not evident. To adequately address a dose-response relationship, the exposure needs to incorporate information regarding frequency of use, duration of use, and intensity of exposure (24). In this study, information was available only on frequency of use. Studies for ovarian cancer have been inconsistent in establishing a dose-response relationship, partly due to the difficulty of accurately quantifying exposure (25).

Our results suggest a possible role of inflammation in the development of endometrial cancer, as talc is a known inflammatory agent (26). Talc may increase endometrial cancer risk by inducing local and/or systemic inflammation. A local inflammatory response would entail activation of macrophages and cytokine production, increased production of reactive oxygen species, increased cell proliferation, DNA damage, and finally, malignant transformation of cells (6). Talc has been shown to produce such responses *in vivo* and *in vitro* (26, 27). The local inflammation mechanism would require talc to reach the uterus. Although few studies are available, there is some evidence of retrograde transport of inert particles through the genital tract (28, 29). Talc particles have been found in human ovarian tissue (12) and human pelvic lymph nodes (30), and an increased risk of ovarian cancer has been noted with talc use before compared with after tubal ligation (22, 23). In addition, talc is a poorly soluble chemical; it is estimated that a 1-μm spherical talc particle in the lung would take 8 years to dissolve (31). These data support the contention that talc particles can migrate and persist in distant organs; furthermore, the uterus is a more accessible site than the ovaries. Exposure of the external genital area to talcum powder may also activate a systemic inflammatory response. Users of talcum powder have lower plasma levels of anti-MUC1 antibodies than nonusers (32). MUC1 is a protein highly expressed by ovarian, breast, and endometrial tumors, and low levels of anti-MUC1 antibodies are associated with poorer prognosis (32, 33). Reducing immunity to MUC-1 may be one mechanism by which talc increases endometrial cancer risk.

Other mechanistic factors that may come into play include chronicity of inflammation (34) and timing of exposure with regard to the phases of the uterine cycle. Any inflammation initiated by genital application of talc

Table 3. Incidence RRs and 95% CIs for talc use and endometrial cancer risk among postmenopausal women in the NHS

	No. of cases	Person-years	Age-adjusted RR (95% CI)	Multivariate RR* (95% CI)
Ever perineal talc use				
No	287	461,381	1.00	1.00
Yes	242	281,958	1.38 (1.16-1.64)	1.21 (1.02-1.44)
Frequency of perineal talc use				
No use	287	461,381	1.00	1.00
Less than once a week	57	76,845	1.22 (0.91-1.62)	1.09 (0.81-1.45)
One to six times a week	87	97,793	1.40 (1.10-1.79)	1.28 (1.00-1.63)
Daily	98	107,320	1.49 (1.18-1.87)	1.24 (0.98-1.56)
P trend†			<0.001	0.04
Regular perineal talc use (at least once a week)				
No	344	538,227	1.00	1.00
Yes	185	205,113	1.40 (1.17-1.68)	1.24 (1.03-1.49)
Sanitary napkin talc use				
No	403	587,317	1.00	1.00
Yes	67	94,233	1.04 (0.80-1.35)	0.98 (0.75-1.27)

*Adjusted for age, parity (0, 1, 2, 3, 4+), age at last birth (nulliparous, <30, 30-34, 35-39, ≥40), age at menarche (≤11, 12, 13, ≥14), age at menopause (<45, 45-49, 50-54, ≥55), OC duration (never, ≤12 mo, 13-36 mo, 37-72 mo, >72 mo), PMH duration (never, past <5 y, past 5+ y, current <5 y, current 5+ y), BMI (continuous), smoking pack-years (0, 1-20, 21-40, 40+), diabetes (no, yes), and family history of uterine cancer (no, yes).

†*P* trend for categories of frequency of perineal talc use weighted as 0, 2, 15.5, and 30 d of use per month.

Published OnlineFirst April 20, 2010; DOI: 10.1158/1055-9965.EPI-09-1221

is likely to be sustained because studies indicate that women start using talcum powder at an early age (35) and continue using it for decades (14). The endometrial tissue is highly proliferative and regenerates with every menstrual cycle. Chronic inflammation following long duration of use of talcum powder may be sufficient to cause carcinogenesis despite the monthly shedding of the endometrial lining. Certain phases of the uterine cycle may also represent windows of particular susceptibility to exposure. For example, exposure during the proliferative phase of the uterine lining may be more likely to cause DNA damage and propagation. On the other hand, the inflammatory response is a natural process in the uterus during the late secretory and menstrual phase. During this period of tissue disintegration, inflammatory cells infiltrate the region, cytokines, prostaglandins, and cyclooxygenase-2 are released, and NF-κB is activated (36). Use of talcum powder during menstruation may interfere with normal immune processes in the uterus and prevent complete shedding of the lining. The inflammation hypothesis as a mechanism for the carcinogenic effects of talc is supported by recent evidence that the risk of ovarian cancer associated with talc is modified by variation in detoxification genes (35), emphasizing that clearance mechanisms are important in reducing risk. Studies show a reduction in risk for endometrial cancer following use of nonsteroidal anti-inflammatory drugs, especially among high-risk individuals (37, 38), supporting the role of inflammation in endometrial cancer.

The strengths of this study include prospectively collected data and adjustment for known risk factors for endometrial cancer, some of which are associated with talc use, such as obesity (39). We adjusted for BMI continuously to reduce confounding, and to minimize potential residual confounding, we secondarily restricted the analysis to normal-weight women. We additionally restricted our analysis to never users of PMH because both BMI and PMH use are strong enough risk factors for endometrial cancer to obscure a modest association with talc, and are linked to chronic systemic inflammation and changes in levels of inflammatory markers (40), respectively. The availability of a single assessment of talc use is a limitation of this study, as this may have resulted in some exposure misclassification during follow-up; however, talc use was assessed when women in the study were above 36 years of age (mean age, 48 years), when never users are unlikely to start using talc, reducing potential exposure misclassification (35). Furthermore, we were unable to explore a dose-response with duration of talc use because information on duration of use was not available.

In summary, we noted a modest positive association between genital use of talcum powder and endometrial cancer risk among postmenopausal women. However, this association needs to be replicated in future and larger studies. Mechanistic studies also are needed to elucidate the process by which talc may increase the risk of carcinogenesis and to provide additional support for this relationship. Studies may include assessing differences in inflammatory markers between talc users and nonusers or *in vitro* studies of the response of endometrial cells to talc particles. In addition, identifying genetically susceptible populations might offer insight into potential mechanisms. Future studies addressing the association between talcum powder use and endometrial cancer risk may provide further evidence for the role of inflammation in endometrial cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

Research grants R01 CA082838 and P01 CA87969 and training grant T32 CA009001 (M. Gates) from the National Cancer Institute, NIH, and the HSPH Cyprus Initiative for the Environment and Public Health funded by the Republic of Cyprus (S. Karageorgi).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 12/04/2009; revised 03/03/2010; accepted 03/05/2010; published OnlineFirst 04/20/2010.

References

1. Huncharek M, Geschwind JF, Kupelnick B. Perineal application of cosmetic talc and risk of invasive epithelial ovarian cancer: a meta analysis of 11,933 subjects from sixteen observational studies. Anticancer Res 2003;23:1955 60.
2. Langseth H, Hankinson SE, Siemiatycki J, Weiderpass E. Perineal use of talc and risk of ovarian cancer. J Epidemiol Community Health 2008;62:358 60.
3. Karageorgi S, Hankinson SE, Kraft P, De Vivo I. Reproductive factors and postmenopausal hormone use in relation to endometrial cancer risk in the Nurses' Health Study cohort 1976 2004. Int J Cancer 2010;126:208 16.
4. Wild P. Lung cancer risk and talc not containing asbestiform fibres: a review of the epidemiological evidence. Occup Environ Med 2006; 63:4 9.
5. Levadie B. Definitions for asbestos and other health related silicates. Philadelphia (PA): ASTM Special Technical Publication; 1984.
6. Baan RA. Carcinogenic hazards from inhaled carbon black, titanium dioxide, and talc not containing asbestos or asbestiform fibers: recent evaluations by an IARC Monographs Working Group. Inhal Toxicol 2007;19 Suppl 1:213 28.
7. Rohl AN, Langer AM, Selikoff IJ, et al. Consumer talcums and powders: mineral and chemical characterization. J Toxicol Environ Health 1976;2:255 84.
8. Harlow BL, Hartge PA. A review of perineal talc exposure and risk of ovarian cancer. Regul Toxicol Pharmacol 1995;21:254 60.
9. Zazenski R, Ashton WH, Briggs D, et al. Talc: occurrence, charac terization, and consumer applications. Regul Toxicol Pharmacol 1995;21:218 29.
10. Merchant JA. Human epidemiology: a review of fiber type and characteristics in the development of malignant and nonmalignant disease. Environ Health Perspect 1990;88:287 93.
11. Pang TWS, Schonfeld FA, Nazar MA. The determination of

tremolite asbestos in talc powder samples. Ann Occup Hyg 1987; 31:219 25.
12. Heller DS, Westhoff C, Gordon RE, Katz N. The relationship between perineal cosmetic talc usage and ovarian talc particle burden. Am J Obstet Gynecol 1996;174:1507 10.
13. Mills PK, Riordan DG, Cress RD, Young HA. Perineal talc exposure and epithelial ovarian cancer risk in the Central Valley of California. Int J Cancer 2004;112:458 64.
14. Wong C, Hempling RE, Piver MS, Natarajan N, Mettlin CJ. Perineal talc exposure and subsequent epithelial ovarian cancer: a case control study. Obstet Gynecol 1999;93:372 6.
15. Rich Edwards JW, Corsano KA, Stampfer MJ. Test of the National Death Index and Equifax Nationwide Death Search. Am J Epidemiol 1994;140:1016 9.
16. Colditz GA, Stampfer MJ, Willett WC, et al. Reproducibility and validity of self reported menopausal status in a prospective cohort study. Am J Epidemiol 1987;126:319 25.
17. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self reported waist and hip circumferences in men and women. Epidemiology 1990;1:466 73.
18. Karageorgi S, Hankinson SE, Kraft P, De Vivo I. Reproductive factors and postmenopausal hormone use in relation to endometrial cancer risk in the Nurses' Health Study cohort 1976 2004. Int J Cancer 2010;126:208 16.
19. Rothman KJ, Poole C. A strengthening programme for weak asso ciations. Int J Epidemiol 1988;17:955 9.
20. Muscat JE, Huncharek MS. Perineal talc use and ovarian cancer: a critical review. Eur J Cancer Prev 2008;17:139 46.
21. Gertig DM, Hunter DJ, Cramer DW, et al. Prospective study of talc use and ovarian cancer. J Natl Cancer Inst 2000;92:249 52.
22. Chang S, Risch HA. Perineal talc exposure and risk of ovarian carci noma. Cancer 1997;79:2396 401.
23. Harlow BL, Cramer DW, Bell DA, Welch WR. Perineal exposure to talc and ovarian cancer risk. Obstet Gynecol 1992;80:19 26.
24. Gustavsson P, Nyberg F, Pershagen G, Scheele P, Jakobsson R, Plato N. Low dose exposure to asbestos and lung cancer: dose response relations and interaction with smoking in a population based case referent study in Stockholm, Sweden. Am J Epidemiol 2002;155:1016 22.
25. Cramer DW, Liberman RF, Titus Ernstoff L, et al. Genital talc expo sure and risk of ovarian cancer. Int J Cancer 1999;81:351 6.
26. Wehner AP. Biological effects of cosmetic talc. Food Chem Toxicol 1994;32:1173 84.
27. Buz'Zard AR, Lau BH. Pycnogenol reduces talc induced neoplastic transformation in human ovarian cell cultures. Phytother Res 2007; 21:579 86.
28. Kunz G, Beil D, Deininger H, Wildt L, Leyendecker G. The dynamics of rapid sperm transport through the female genital tract: evidence from vaginal sonography of uterine peristalsis and hysterosalpingos cintigraphy. Hum Reprod 1996;11:627 32.
29. Egli GE, Newton M. The transport of carbon particles in the human female reproductive tract. Fertil Steril 1961;12:151 5.
30. Cramer DW, Welch WR, Berkowitz RS, Godleski JJ. Presence of talc in pelvic lymph nodes of a woman with ovarian cancer and long term genital exposure to cosmetic talc. Obstet Gynecol 2007;110:498 501.
31. Jurinski JB, Rimstidt JD. Biodurability of talc. Am Mineral 2001;86: 392 9.
32. Cramer DW, Titus Ernstoff L, McKolanis JR, et al. Conditions asso ciated with antibodies against the tumor associated antigen MUC1 and their relationship to risk for ovarian cancer. Cancer Epidemiol Biomarkers Prev 2005;14:1125 31.
33. Sivridis E, Giatromanolaki A, Koukourakis MI, Georgiou L, Anastasiadis P. Patterns of episialin/MUC1 expression in endome trial carcinomas and prognostic relevance. Histopathology 2002; 40:92 100.
34. Hussain SP, Harris CC. Inflammation and cancer: an ancient link with novel potentials. Int J Cancer 2007;121:2373 80.
35. Gates MA, Tworoger SS, Terry KL, et al. Talc use, variants of the GSTM1, GSTT1, and NAT2 genes, and risk of epithelial ovarian can cer. Cancer Epidemiol Biomarkers Prev 2008;17:2436 44.
36. Modugno F, Ness RB, Chen C, Weiss NS. Inflammation and endo metrial cancer: a hypothesis. Cancer Epidemiol Biomarkers Prev 2005;14:2840 7.
37. Moysich KB, Baker JA, Rodabaugh KJ, Villella JA. Regular analgesic use and risk of endometrial cancer. Cancer Epidemiol Biomarkers Prev 2005;14:2923 8.
38. Viswanathan AN, Feskanich D, Schernhammer ES, Hankinson SE. Aspirin, NSAID, and acetaminophen use and the risk of endometrial cancer. Cancer Res 2008;68:2507 13.
39. Rosenblatt KA, Mathews WA, Daling JR, Voigt LF, Malone K. Characteristics of women who use perineal powders. Obstet Gynecol 1998;92:753 6.
40. Miller AP, Chen YF, Xing D, Feng W, Oparil S. Hormone replacement therapy and inflammation: interactions in cardiovascular disease. Hypertension 2003;42:657 63.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Perineal Use of Talcum Powder and Endometrial Cancer Risk

Stalo Karageorgi, Margaret A. Gates, Susan E. Hankinson, et al.

Cancer Epidemiol Biomarkers Prev 2010;19:1269-1275. Published OnlineFirst April 20, 2010.

Exhibit 113

NIH NATIONAL CANCER INSTITUTE

Ovarian, Fallopian Tube, and Primary Peritoneal Cancer Prevention (PDQ®)–Health Professional Version

Go to Patient Version

Who Is at Risk?

Ovarian cancer is a rare disease, with carcinomas comprising approximately 90% of tumors and germ cell and stromal tumors accounting for the remainder. Ovarian carcinoma is a disease that predominantly affects postmenopausal women. Ovarian carcinomas consist of several histopathologic types, with high-grade serous being both the most common and most lethal. The category of ovarian borderline tumor or tumor of low-malignant potential, which historically had been considered in the context of ovarian cancer, is now generally considered a nonmalignant entity, although it has a postulated relationship with the development of some histologic subtypes of low-grade ovarian carcinomas.[1]

Risk factors for ovarian cancer include a family history of breast and/or ovarian cancer and inheritance of deleterious mutations in *BRCA1*, *BRCA2*, and selected other high-penetrance genes.[2-6] (Refer to the PDQ summary on Genetics of Breast and Gynecologic Cancers for more information.) Other risk factors for ovarian cancer include obesity, tall height, endometriosis, and the use of postmenopausal hormone therapy.[7-9]

Associations of some risk factors with ovarian cancer vary by histopathologic subtype. The association of endometriosis with ovarian cancer is stronger for nonserous subtypes, especially clear cell carcinoma and endometrioid subtypes.[10] Further, among carriers of deleterious mutations in *BRCA1* or *BRCA2*, increasing evidence suggests that many tumors previously classified as ovarian high-grade serous carcinoma may develop from malignant cells arising in the tubal epithelium (serous tubal intraepithelial carcinoma [STIC]), although these tumors continue to be referred to as *ovarian* cancers in most writings. It is hypothesized that high-grade serous carcinomas among individuals who are not carriers of mutations in *BRCA1* or *BRCA2* may also develop in the fallopian tube, but few STICs have been identified among these women in the absence of concurrent high-stage disease. Further, data suggest that the distinction of high-grade serous carcinomas from other histologic types of high-grade carcinomas, particularly endometrioid carcinomas, is not reliable. Reported rates of mucinous carcinoma diagnoses have declined dramatically, but expert pathology reviews suggest that this reflects increased recognition of metastases from occult gastrointestinal primary tumors to the ovary, rather than a true decline in rates of ovarian primary tumors.[11]

Factors associated with a decreased risk of ovarian cancer include multiparity, use of oral contraceptives, multiple pregnancies, breastfeeding, tubal ligation, and salpingectomy.[12-15] Compared with nulliparous women, the risk of ovarian cancer is reduced by 30% to 60% among parous women, with additive protection for each additional birth.[16,17]

References

1. Kurman RJ, Carcangiu ML, Young RH, eds.: WHO Classification of Tumours of Female Reproductive Organs. 4th ed. Lyon, France: International Agency for Research on Cancer, 2014.

2. Bolton KL, Ganda C, Berchuck A, et al.: Role of common genetic variants in ovarian cancer susceptibility and outcome: progress to date from the Ovarian Cancer Association Consortium (OCAC). J Intern Med 271 (4): 366-78, 2012. [PUBMED Abstract]

3. Weissman SM, Weiss SM, Newlin AC: Genetic testing by cancer site: ovary. Cancer J 18 (4): 320-7, 2012 Jul-Aug. [PUBMED Abstract]

4. Hunn J, Rodriguez GC: Ovarian cancer: etiology, risk factors, and epidemiology. Clin Obstet Gynecol 55 (1): 3-23, 2012. [PUBMED Abstract]

5. Pal T, Akbari MR, Sun P, et al.: Frequency of mutations in mismatch repair genes in a population-based study of women with ovarian cancer. Br J Cancer 107 (10): 1783-90, 2012. [PUBMED Abstract]

6. Gayther SA, Pharoah PD: The inherited genetics of ovarian and endometrial cancer. Curr Opin Genet Dev 20 (3): 231-8, 2010. [PUBMED Abstract]

7. Lacey JV Jr, Brinton LA, Leitzmann MF, et al.: Menopausal hormone therapy and ovarian cancer risk in the National Institutes of Health-AARP Diet and Health Study Cohort. J Natl Cancer Inst 98 (19): 1397-405, 2006. [PUBMED Abstract]

8. Trabert B, Wentzensen N, Yang HP, et al.: Ovarian cancer and menopausal hormone therapy in the NIH-AARP diet and health study. Br J Cancer 107 (7): 1181-7, 2012. [PUBMED Abstract]

9. Lahmann PH, Cust AE, Friedenreich CM, et al.: Anthropometric measures and epithelial ovarian cancer risk in the European Prospective Investigation into Cancer and Nutrition. Int J Cancer 126 (10): 2404-15, 2010. [PUBMED Abstract]

10. Poole EM, Lin WT, Kvaskoff M, et al.: Endometriosis and risk of ovarian and endometrial cancers in a large prospective cohort of U.S. nurses. Cancer Causes Control 28 (5): 437-445, 2017. [PUBMED Abstract]

11. Seidman JD, Kurman RJ, Ronnett BM: Primary and metastatic mucinous adenocarcinomas in the ovaries: incidence in routine practice with a new approach to improve intraoperative diagnosis. Am J Surg Pathol 27 (7): 985-93, 2003. [PUBMED Abstract]

12. Garg PP, Kerlikowske K, Subak L, et al.: Hormone replacement therapy and the risk of epithelial ovarian carcinoma: a meta-analysis. Obstet Gynecol 92 (3): 472-9, 1998. [PUBMED Abstract]

13. Lacey JV Jr, Mink PJ, Lubin JH, et al.: Menopausal hormone replacement therapy and risk of ovarian cancer. JAMA 288 (3): 334-41, 2002. [PUBMED Abstract]

14. Mills PK, Riordan DG, Cress RD, et al.: Hormone replacement therapy and invasive and borderline epithelial ovarian cancer risk. Cancer Detect Prev 29 (2): 124-32, 2005. [PUBMED Abstract]

15. Calle EE, Rodriguez C, Walker-Thurmond K, et al.: Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 348 (17): 1625-38, 2003. [PUBMED Abstract]

16. Permuth-Wey J, Sellers TA: Epidemiology of ovarian cancer. Methods Mol Biol 472: 413-37, 2009. [PUBMED Abstract]

17. Wentzensen N, Poole EM, Trabert B, et al.: Ovarian Cancer Risk Factors by Histologic Subtype: An Analysis From the Ovarian Cancer Cohort Consortium. J Clin Oncol 34 (24): 2888-98, 2016. [PUBMED Abstract]

Overview

Note: Separate PDQ summaries on Ovarian, Fallopian Tube, and Primary Peritoneal Cancer Screening and Ovarian Epithelial, Fallopian Tube, and Primary Peritoneal Cancer Treatment are also available.

Factors With Adequate Evidence of an Increased Risk of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Family history and inherited susceptibility to ovarian, fallopian tube, and primary peritoneal cancer

Based on solid evidence, women with a family history of ovarian cancer, especially in a first-degree relative, and those with an inherited predisposition to ovarian cancer, such as a *BRCA1* or *BRCA2* mutation, have an increased risk of developing ovarian cancer. (Refer to the PDQ summary on Genetics of Breast and Gynecologic Cancers for more information.)

Endometriosis

Based on fair evidence, self-reported and laparoscopically confirmed endometriosis is associated with an increased risk of ovarian cancer.[1,2] The association is stronger with nonserous histologic subtypes, specifically endometrioid and clear cell carcinomas.[2,3]

Magnitude of Effect: Modest with observed relative risks (RRs) of 1.8 to 2.4.

Study Design Cohort and case control studies

Internal Validity Good

Consistency Fair

External Validity Good

Hormone replacement therapy

Based on fair evidence, current or recent hormone therapy is associated with a small increased risk of ovarian cancer. Risks attenuate after hormone therapy is discontinued. Risks did not differ by preparation type (estrogen only vs. combined estrogen/progestin).[4,5]

Magnitude of Effect: Modest with observed RRs of 1.20 to 1.8.

Study Design: One randomized clinical trial, cohort and case-control studies.

Internal Validity: Good.

Consistency: Fair.

External Validity: Good.

Obesity and height

Based on fair evidence, increases in height and body mass index (BMI) are associated with a modest increased risk of ovarian cancer.

Magnitude of Effect: Based on an overview analysis of 25,157 women with ovarian cancer and 81,211 women without ovarian cancer from 47 epidemiological studies, the RR of ovarian cancer per 5 cm increase in height is 1.07 (95% confidence interval [CI], 1.05–1.09). The RR of ovarian cancer per 5 kg/m^2 increase in BMI is 1.10 (95%

CI, 1.07–1.13) among never-users of hormone therapy and 0.95 (95% CI, 0.92–0.99) among ever-users of hormone therapy.[6]

Study Design: Cohort and case-control studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Factors With Adequate Evidence of a Decreased Risk of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Oral contraceptives: benefits

Based on solid evidence, oral contraceptive use is associated with a decreased risk of developing ovarian cancer.

Magnitude of Effect: The degree of risk reduction varies by duration of oral contraceptive use and time since last use. For 1 to 4 years of oral contraceptive use, the RR reduction is 22%, and for 15 or more years of use, the RR reduction is 56%. The reduction in risk persisted for more than 30 years after use was discontinued, but the degree of reduction attenuated over time. The risk reduction per 5 years of oral contraceptive use was 29% for women who discontinued use less than 10 years ago and decreased to 15% for women who discontinued use 20 to 29 years ago. Ten years of use reduced cancer incidence before age 75 years from 1.2 to 0.8 per 100 users and reduced mortality from 0.7 to 0.5 per 100 users. The number needed-to-treat for 5 years was estimated to be about 185 women.

Study Design: Multiple case-control and cohort studies; meta-analyses.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Oral contraceptives: harms

Based on solid evidence, combined current use of estrogen-progestin oral contraceptive use is associated with an increased risk of venous thromboembolism, particularly among smokers, for whom use is contraindicated. Oral contraceptives are not associated with a long-term increased risk of breast cancer but may be associated with a short-term increased risk while a woman is taking oral contraceptives. The risk of breast cancer declines with time since last use.

Magnitude of Effect: The risks may vary by preparation. Overall, the absolute risk of venous thromboembolism is about three events per 10,000 women per year while taking oral contraceptives. The risk is modified by smoking. Breast cancer risk among long-term (>10 years) current users is estimated at one extra case per year per 100,000 women. The risk dissipates with time since last use.

Study Design: Observational studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Tubal ligation: benefits

Based on solid evidence, tubal ligation is associated with a decreased risk of ovarian cancer.

Magnitude of Effect: Adjusting for other forms of contraception, tubal ligation provides a relative reduction in the odds of developing ovarian cancer of about 30%.

Study Design: Multiple case-control studies and cohort studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Tubal ligation: harms

Based on fair evidence, harms include surgical risks, including the following:[7]

- Major morbidity including blood transfusion, reoperation, or hospital readmission (rate of 1.0 per 100 procedures).
- Minor morbidity including postoperative fever, urinary tract infections, or wound infections (rate of 6.0 per 100 procedures).

Multiparity

Based on good evidence, multiparity is associated with a decreased risk of ovarian cancer.

Magnitude of Effect: Based on good evidence from multiple observational epidemiological studies, parous women have an approximately 30% lower ovarian cancer risk than nulliparous women.[6,8,9]

Study Design: Observational epidemiologic studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Salpingectomy

Based on limited data, salpingectomy is associated with a decrease in risk of ovarian cancer.

Magnitude of Effect: Approximately 50% decrease for bilateral salpingectomy, less protection for unilateral salpingectomy.

Study Design: Observational epidemiologic studies from several different countries.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Breastfeeding

Based on solid evidence, breastfeeding is associated with a decreased risk of ovarian cancer.

Magnitude of Effect: 2% decrease with every month of breastfeeding.[10]

Study Design: Multiple case-control and cohort studies; meta-analysis.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Risk reducing bilateral salpingo oophorectomy: benefits

Based on solid evidence, risk-reducing bilateral salpingo-oophorectomy is associated with a decreased risk of ovarian cancer. Peritoneal carcinomatosis has been reported rarely following surgery. Risk-reducing surgery is generally reserved for women at high risk of developing ovarian cancer, such as women who have an inherited susceptibility to ovarian cancer

Magnitude of Effect: 90% reduction in risk of ovarian cancer observed among women with a *BRCA1* or *BRCA2* mutation

Study Design: Multiple case-control studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Risk-reducing bilateral salpingo-oophorectomy: harms

Based on solid evidence, prophylactic oophorectomy among women who are still menstruating at the time of surgery is associated with infertility, vasomotor symptoms, decreased sexual interest, vaginal dryness, urinary frequency, decreased bone-mineral density, and increased cardiovascular disease.

Magnitude of Effect: Reported prevalence of vasomotor symptoms varies from 41% to 61.4% among women who underwent oophorectomy before natural menopause Women with bilateral oophorectomy who did not take hormone therapy were twice as likely to have moderate or severe hot flashes compared with women who underwent natural menopause. The RR of cardiovascular disease among women with bilateral oophorectomy and early menopause was 4.55 (95% CI, 2.56–9.01).

Study Design: Cohort and case-control studies.

Internal Validity: Good.

Consistency: Good.

External Validity: Good.

Areas of Uncertainty

Ovarian hyperstimulation for infertility treatment

Evidence is poor to determine the association between ovarian hyperstimulation and the risk of ovarian cancer. Risk of ovarian cancer may be increased among women who remain nulligravid after being treated with ovarian stimulating medications.

Magnitude of Effect: Uncertain—risk of invasive ovarian cancer may be increased among women who remain nulligravid after treatment; risk of borderline ovarian tumors may be increased among women treated with infertility drugs.

Study Design: Cohort and case-control studies; systematic review.

Internal Validity: Fair.

Consistency: Poor.

External Validity: Fair.

References

1. Poole EM, Lin WT, Kvaskoff M, et al.: Endometriosis and risk of ovarian and endometrial cancers in a large prospective cohort of U.S. nurses. Cancer Causes Control 28 (5): 437-445, 2017. [PUBMED Abstract]
2. Pearce CL, Templeman C, Rossing MA, et al.: Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. Lancet Oncol 13 (4): 385-94, 2012. [PUBMED Abstract]
3. Mogensen JB, Kjær SK, Mellemkjær L, et al.: Endometriosis and risks for ovarian, endometrial and breast cancers: A nationwide cohort study. Gynecol Oncol 143 (1): 87-92, 2016. [PUBMED Abstract]
4. Mørch LS, Løkkegaard E, Andreasen AH, et al.: Hormone therapy and ovarian cancer. JAMA 302 (3): 298-305, 2009. [PUBMED Abstract]
5. Beral V, Gaitskell K, Hermon C, et al.: Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. Lancet 385 (9980): 1835-42, 2015. [PUBMED Abstract]
6. Braem MG, Onland-Moret NC, van den Brandt PA, et al.: Reproductive and hormonal factors in association with ovarian cancer in the Netherlands cohort study. Am J Epidemiol 172 (10): 1181-9, 2010. [PUBMED Abstract]
7. Lawrie TA, Kulier R, Nardin JM: Techniques for the interruption of tubal patency for female sterilisation. Cochrane Database Syst Rev (9): CD003034, 2015. [PUBMED Abstract]
8. Fortner RT, Ose J, Merritt MA, et al.: Reproductive and hormone-related risk factors for epithelial ovarian cancer by histologic pathways, invasiveness and histologic subtypes: Results from the EPIC cohort. Int J Cancer 137 (5): 1196-208, 2015. [PUBMED Abstract]
9. Yang HP, Trabert B, Murphy MA, et al.: Ovarian cancer risk factors by histologic subtypes in the NIH-AARP Diet and Health Study. Int J Cancer 131 (4): 938-48, 2012. [PUBMED Abstract]
10. Feng LP, Chen HL, Shen MY: Breastfeeding and the risk of ovarian cancer: a meta-analysis. J Midwifery Womens Health 59 (4): 428-37, 2014 Jul-Aug. [PUBMED Abstract]

Description of the Evidence

Incidence and Mortality

In 2019, it is estimated that 22,530 new cases of ovarian cancer will be diagnosed and 13,980 deaths due to ovarian cancer will occur.[1] Incidence and mortality rates are higher among whites than among blacks, but

statistically significant decreases in incidence and mortality rates have been observed among both whites and blacks.[2] In 2014, the overall incidence rate for ovarian carcinoma among women aged 65 years and older was 41.9 cases per 100,000 women-years.[3] Given that the Surveillance, Epidemiology, and End Results Program does not adjust for oophorectomy or salpingectomy, racial differences in the prevalence of women who had undergone these procedures could bias racial rate comparisons. A statistically significant decrease in delayed adjusted incidence of 0.9% among whites from 1987 to 2012 and 0.2% among blacks from 1992 to 2012 was observed. A statistically significant decrease in mortality rates of 2.0% per year among whites from 2002 to 2012 and 1.3% per year among blacks from 1992 to 2012 was observed. The population lifetime risk of ovarian cancer is 1.3%; the population lifetime risk of dying from ovarian cancer is 0.97%.[2]

Histology and Pathogenesis of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Ovarian carcinoma is a biologically and clinically heterogeneous class of tumors that includes several major subtypes: serous, mucinous, endometrioid, and clear cell. Classification of ovarian carcinomas into type I and type II tumors has been proposed. In this system, type I tumors include the following:[4]

1. Endometriosis-related subtypes, such as endometrioid, clear cell, and seromucinous.
2. Low-grade serous.
3. Mucinous and malignant Brenner tumors.

Among type I tumors, endometrioid and clear cell carcinomas are numerically predominant and most important clinically. In general, type I ovarian carcinomas present at a lower stage than type II tumors and portend a better prognosis.

Type II tumors are comprised mainly of high-grade serous carcinomas, the most common and lethal of all ovarian carcinoma subtypes. These cancers usually present with symptomatic bulky stage III or IV disease and ascites. Many, but possibly not all, high-grade serous carcinomas appear to arise from malignant *in situ* lesions in the epithelium of the fallopian tube fimbria, which spread to the ovaries secondarily, but continue to be referred to as ovarian carcinomas. Evidence for a tubal origin is based mainly on examination of risk-reducing salpingo-oophorectomy specimens, performed among *BRCA1/BRCA2* mutation carriers, in which incidental low-volume disease enables recognition of serous tubal intraepithelial carcinoma (STIC). However, not all women with high-grade serous carcinomas have identifiable STIC and few studies of the fallopian tubes among women who are not carriers of *BRCA1/BRCA2* mutations have been performed, suggesting that pathogenesis of these tumors is not fully known. Serous carcinomas can be further divided on the basis of molecular characteristics.[5]

The heterogeneity in the etiology and pathogenesis of different ovarian cancer subtypes and variability in the classification of tumors over time and between studies pose challenges for interpretation of etiologic data. Ovarian cancer is a rare cancer, thus sample size and power of studies to detect moderate associations by cancer subtype is limited. However, clearer subtyping of cancers may assist in improving our understanding of the etiology of ovarian malignancies in future studies.

Factors With Adequate Evidence of an Increased Risk of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Family history and inherited susceptibility to ovarian, fallopian tube, and primary peritoneal cancer

Some women are at an increased risk because of an inherited mutation, with the magnitude of that risk dependent on the affected gene and specific mutation. Underlying ovarian cancer risk can be assessed through accurate pedigrees and/or genetic markers of risk. Because of uncertainties about cancer risks associated with certain specific gene mutations, genetic information may be difficult to interpret outside of families with a high incidence of ovarian cancer.

This summary does not address multiple genetic syndromes or women who are at high risk because of inherited genetic factors. (Refer to the PDQ summaries on Genetics of Breast and Gynecologic Cancers and Genetics of Colorectal Cancer for specific information related to ovarian cancer risk associated with multiple genetic syndromes and ovarian cancer in *BRCA1/BRCA2* mutation carriers.)

Hormone replacement therapy/hormone therapy

A meta-analysis of 52 studies (17 prospective and 35 retrospective) including 21,488 ovarian cancers found increased risks with current or recent hormone replacement use in prospective studies (relative risk [RR], 1.37; 95% confidence interval [CI], 1.29–1.46), with similar results for retrospective designs. Significant relationships were found for serous and endometrioid subtypes.[6] Recent use was strongly related to risk even among women who had used hormone replacement for less than 5 years (RR, 1.41; 95% CI, 1.32–1.50). Risk declined among women who had discontinued use, with greater effects for longer periods of cessation. Risks did not differ by preparation types (estrogen only vs. combined estrogen/progestin). Risks also did not differ by age at use.[7,8]

Obesity and height

Ovarian cancer risk increases with increasing height and weight (body mass index [BMI]).[9] The Collaborative Group on Epidemiological Studies of Ovarian Cancer compiled individual data, both published and unpublished, from 47 epidemiological studies including 12,157 women with ovarian cancer and 81,311 controls. RR increased significantly with increasing height (1.07 per 5 cm height) and with increasing BMI (1.10 per 5 kg/m^2). These findings were unaffected by other factors known to be associated with ovarian cancer risk, with the exception that ever-users of hormone therapy had no increased risk with increasing BMI. Given that height, weight, and BMI are thought to be strongly correlated, separating out the individual effects can be difficult. Ovarian cancer mortality has also been shown to be increased in obese women.[10,11]

Factors With Adequate Evidence of a Decreased Risk of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Oral contraceptives

A collaborative analysis was performed of individual data from 23,257 women with ovarian cancer and 87,303 women without ovarian cancer from 45 studies in 21 countries.[12] The studies included 13 prospective studies, 19 population-based case-control studies, and 12 hospital-based case-control studies. Oral contraceptive use was associated with a dose-response effect by duration of use, without observed changes in risk reduction by decade of use from the 1960s to 1980s, over which time the amount of estrogen in oral contraceptives was approximately halved. No risk reduction was observed for women who used oral contraceptives for less than 1 year. The risk reduction associated with use from 1 to 4 years, 5 to 9 years, 10 to 14 years, and 15 years or more was 0.78 (99% CI, 0.73–0.893), 0.64 (99% CI, 0.59–0.69), 0.56 (99% CI, 0.50–0.62), and 0.42 (99% CI, 0.36–0.49), respectively. The observed risk reduction persisted after cessation of oral contraceptive therapy but attenuated over time since last use. The proportional reduction in risk per 5 years of use was 29% (95% CI, 23%–34%) for

women who had discontinued use within the last 10 years; the reduction in risk was 15% (95% CI, 9%–21%) for women who discontinued use 20 to 29 years ago.

A meta-analysis, in which the primary analysis was restricted to 24 case-control and cohort studies published since 2000 to reflect more recent types of oral contraceptive preparations, also observed a dose-response by duration of use.[13] The risk reduction among women using oral contraceptives for more than 1 year but less than 5 years was 0.77 (95% CI, 0.66–0.89), and for women using oral contraceptives for more than 10 years, the risk reduction was 0.43 (95% CI, 0.37–0.51). The authors estimated that 185 women needed to be treated for 5 years to prevent one case of ovarian cancer. Based on an estimated lifetime risk of 1.38% and prevalence of ever-use of oral contraceptives of 83%, the authors estimated a lifetime reduction of ovarian cancer attributable to oral contraceptives of 0.54%.

(Refer to the PDQ summary on Genetics of Breast and Gynecologic Cancers for specific information related to ovarian cancer risk among *BRCA1/BRCA2* mutation carriers.)

Depot-medroxyprogesterone acetate

Limited information is available on the use of injectable progestational contraceptives (depot-medroxyprogesterone acetate [DMPA]) and the risk of ovarian cancer; studies are confounded by the use of other contraceptive methods, particularly oral contraceptives. A hospital-based study conducted in Mexico and Thailand, with 224 cases and 1,781 controls (the World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives), did not observe an association between DMPA and ovarian cancer (RR, 1.07; 95% CI, 0.6–1.8).[14] However, only 22 of the cases had ever used DMPA and nine of these had used it for 6 months or less.

A subsequent multicenter study conducted in 12 hospitals in Thailand, including 330 cases and 982 matched controls, observed a statistically significant decreased risk of ovarian cancer associated with DMPA use, controlling for oral contraceptive use and other associated factors (odds ratio [OR], 0.52; 95% CI, 0.33–0.88). A dose-response association was observed but the sample size was limited in longer-term use categories.[15]

Tubal ligation

A meta-analysis of 16 case-control studies, three retrospective studies, and two prospective cohort studies observed a decreased risk of ovarian cancer associated with tubal ligation (RR, 0.66; 95% CI, 0.60–0.73).[16] The reduced risk was observed up to 14 years after tubal ligation. A population-based case-control study of 902 cases and 1,802 controls published subsequent to the meta-analysis observed an adjusted OR of 0.62 (95% CI, 0.51–0.75) associated with a history of a tubal ligation.[17] The association was adjusted for oral contraceptive use, which was also associated with a lower risk of ovarian cancer (OR, 0.62; 95% CI, 0.47–0.85) and other risk factors.[17]

Another pooling project with primary data from 13 population-based case-control studies examined the association between tubal ligation and ovarian cancer risk and included 7,942 epithelial ovarian cancers, and 13,904 controls.[18] Overall, tubal ligation was associated with a 29% reduction in risk (OR, 0.71; 95% CI, 0.66–0.77). The observed risk reduction varied by subtype of invasive cancers and was 52% (OR, 0.48; 95% CI, 0.40–49) for endometrioid cancer; 48% (OR, 0.52; 95% CI, 0.40–0.67) for clear cell cancer; 32% (OR, 0.68; 95% CI, 0.52–89) for mucinous cancer; and 19% (OR, 0.81; 95% CI, 0.74–0,89) for serous cancer.

A pooled analysis from 21 prospective cohort studies examined 14 hormonal, reproductive, and lifestyle factors by histologic subtype among 5,584 invasive ovarian cancers within a total sample of 1.3 million women. Overall, tubal ligation was associated with an 18% reduction in risk (OR, 0.82; 95% CI, 0.73–0.93). The observed risk

reduction varied by subtype of invasive cancer and was 40% (OR, 0.60; 95% CI, 0.41–88) for endometrioid cancer; 65% (OR, 0.35; 95% CI, 0.18–0.69) for clear cell cancer; and 9% (OR, 0.91; 95% CI, 0.79–1.06) for serous cancer. There was a nonsignificant increase in risk of 1% (OR, 1 01; 95% CI, 0 60 1 71) for mucinous cancer [19]

Breastfeeding

A meta-analysis [20] that included five prospective studies and 30 case-control studies examined the association between breastfeeding and the risk of ovarian cancer. Any breastfeeding was associated with a decreased risk of ovarian cancer (RR, 0 76; 95% CI, 0 69 0 83) The risk of ovarian cancer decreased 8% for every 5 month increase in duration of breastfeeding (95% CI, 0.90–0.95). Another meta-analysis that included five prospective studies and 35 case-control studies found that any breastfeeding was associated with a decreased risk of ovarian cancer (RR, 0.70; 95% CI, 0.64–0.76). These results are consistent with a previous meta-analysis and further support the prior finding of a suggested association between increased duration of breastfeeding and greater levels of protection.[21] Another meta-analysis of 19 studies, including four cohort and 15 case-control studies found an overall decreased risk of ovarian cancer with an OR of 0.66 (95% CI, 0.57–0.76) and an association with duration (2% decrease per month). The benefit of breastfeeding was greatest for the first 8 to 10 months.[22]

Risk-reducing salpingo-oophorectomy

Risk-reducing surgery is an option considered by women who are at high risk of ovarian cancer, such as those with an inherited susceptibility to cancer. (Refer to the Oral contraceptives section in the PDQ summary on Genetics of Breast and Gynecologic Cancers for more information on this as a risk reducing intervention) Among women in the general population, opportunistic salpingectomy, oophorectomy, or salpingo-oophorectomy have been considered as possible interventions at the time of surgery for other benign indications. Salpingectomy has also been discussed as a preferred means of sterilization.[23,24]

Harms

Risks associated with benign oophorectomy (with or without salpingectomy or hysterectomy) have been analyzed in six published studies. Studies of three cohorts found that oophorectomy performed before menopause (age 45 or 50 years) was associated with increased overall mortality, likely related to cardiovascular disease. This finding was noted particularly among individuals not using hormone replacement. In the Women's Health Initiative, bilateral salpingo-oophorectomy was not associated with increased mortality. In the National Health and Nutrition Examination Survey (NHANES III), oophorectomy overall was not related to mortality, but mortality was increased among obese women younger than 40 years who did not use hormone replacement The California Teachers Study did not find a mortality risk with oophorectomy, but only 3% of women did not use hormone replacement. Overall, data suggest that oophorectomy among younger women likely increases overall mortality and that this risk may be attenuated with hormone replacement.[25-30]

Salpingectomy

Data relating salpingectomy to risk of ovarian/tubal cancer are limited, but consistent. A meta-analysis of three studies found an OR of 0 51 (95% CI, 0 35 0 71) for risk of these cancers among women who had undergone salpingectomy, compared with women who had intact fallopian tubes.[31] These studies included a Swedish record linkage study conducted from 1973 to 2009 with a mean follow-up of 23 years, which found the following hazard ratios (HRs) for risk of ovarian cancer compared with women who had not undergone surgery:

- For hysterectomy, the HR was 0.79 (95% CI, 0.70–0.88).
- For hysterectomy with bilateral salpingo-oophorectomy, the HR was 0.06 (95% CI, 0.03–0.12).

- For salpingectomy, the HR was 0.65 (95% CI, 0.52–0.81).
- For sterilization procedures, the HR was 0.72 (95% CI, 0.64–0.81).

Protection for bilateral salpingectomy was approximately twice that for unilateral salpingectomy.[32] This report included limited covariate data but results were similar to other smaller studies included in the meta-analysis.

Limited data based on circulating surrogate markers of ovarian reserve suggest that salpingectomy does not have an adverse effect on ovarian function.[33,34]

Factors With Inadequate Evidence of an Association Risk of Ovarian, Fallopian Tube, and Primary Peritoneal Cancer

Dietary factors

No consistent association has been observed between a variety of dietary factors and the risk of ovarian cancer.

A systematic review and meta-analysis that included 23 case-control studies and three cohort studies found no evidence of an association between alcohol use and epithelial ovarian cancer.[35]

A case-control study of the Healthy Eating Index (HEI), based on current U.S. Department of Agriculture dietary guidelines, found no association between the highest HEI score and ovarian cancer risk for any specific food group.[36] A systematic review of the role of diet in ovarian cancer included only prospective studies, with at least 200 reported cases in the publications.[37] Twenty-four publications from ten cohort studies were reviewed and no dietary factors were consistently associated with the risk of ovarian cancer.

Aspirin and nonsteroidal anti-inflammatory drugs

A systematic review and meta-analysis of 21 observational studies found a decreased risk of invasive ovarian cancer associated with aspirin use (RR, 0.88; 95% CI, 0.79–0.98), but no statistically significant association with nonsteroidal anti-inflammatory drugs (NSAIDs).[38] A study published subsequent to that review examined NSAID use and ovarian cancer risk in the National Institutes of Health-AARP Diet and Health Study. No association was observed between the development of ovarian cancer and regular aspirin use (RR, 1.06; 95% CI, 0.87–1.29) or NSAID use (RR, 0.93; 95% CI, 0.74–1.15).[39] A population-based case-control study [40] of 902 incident cases and 1,802 population controls observed a decreased risk of ovarian cancer associated with continual use (0.71; 95% CI, 0.53–0.97) or low-dose daily use (0.72; 95% CI, 0.53–0.97). In that study, selective cyclo-oxygenase-2 NSAIDs but not nonselective NSAIDs were associated with a decreased risk of ovarian cancer (OR, 0.60; 95% CI, 0.39–0.94). A cohort analysis of about 200,000 women in the Nurses' Health Studies, which used detailed data about the intensity and duration of aspirin use over time, showed a reduced HR for ovarian cancer of 0.77 (95% CI, 0.61–0.96) for low-dose aspirin use (≤100 mg/d) but no reduction for standard-dose aspirin use (HR, 1.17; 95% CI, 0.92–1.49).[41]

Perineal talc exposure

The weight of evidence does not support an association between perineal talc exposure and an increased risk of ovarian cancer. Results from case-control and cohort studies are inconsistent. A meta-analysis of 16 studies observed an increased risk with the use of talc (RR, 1.33; 95% CI, 1.16–1.45); however, a dose response relationship was not found.[42] A pooled analysis from the Ovarian Cancer Association Consortium, composed of multiple case-control studies, included 8,525 cases and 9,859 controls, found a modest increased risk of epithelial ovarian cancer associated with genital powder use (OR, 1.24; 95% CI, 1.15–1.33), but the trend across

increasing lifetime number of applications was not statistically significant (*P* trend = .17).[43] A population-based case-control study of African American women in the United States found an association between genital powder use and risk of epithelial ovarian cancer (OR, 1.44; 95% CI, 1.11–1.86).[44] In this study of 584 cases and 745 controls, a dose-response relationship for *any* genital powder use was reported. Specifically, among *any* genital powder use, daily powder use was associated with increased adjusted OR of developing ovarian cancer (OR, 1.71; 95% CI, 1.26–2.33) compared with less than daily use (OR, 1.12; 95% CI, 0.80–1.58). A cohort study among nurses did not observe a risk of ovarian cancer associated with perineal talc use (RR, 1.09; 95% CI, 0.86–1.37) and there was no evidence of increased risk with increasing frequency of use.[45] Another prospective study, The Women's Health Initiative, examined the association between perineal powder use and the development of ovarian cancer among 61,576 women without a history of cancer at enrollment and who provided exposure information. Among this group, 429 cases of ovarian cancer occurred. Powder use on genitals, sanitary napkins, and diaphragms was examined individually and as a combined exposure. Women were followed for a mean of 12.4 years. An association of ovarian cancer with ever-use was not found when analyzed either by individual method of exposure or by overall combined exposure. The observed risk (hazard ratio) for combined exposure to perineal powder was 1.06 (95% CI, 0.87–1.28) and there was no increased risk observed for increasing duration of use.[46]

Areas of Uncertainty

Ovarian hyperstimulation due to infertility treatment

Controversy persists concerning the association between ovarian hyperstimulation and ovarian cancer. Results of a systematic review and meta-analysis of nine cohort studies comprised 109,969 women who were exposed to ovarian hyperstimulation for infertility treatment (i.e., *in vitro* fertilization [IVF]), with 76 incident ovarian cancer cases observed, provided inconclusive evidence for an association.[47] An increased risk of ovarian cancer was observed when the comparison group was the general population (RR, 1.50; 95% CI, 1.17–1.92), but no statistically significant increased risk was observed when the reference group was unexposed infertile women (RR, 1.26; 95% CI, 0.62–2.55). A major limitation was that only one of the cohort studies included in the meta-analysis had a follow-up period longer than 10 years for those exposed to IVF.

A Cochrane systematic review that included 11 case-control studies and 14 cohort studies, for a total of 186,972 women, was also indeterminate for an association. Summary statistics were not calculated because of methodological and clinical heterogeneity. Among seven cohort studies that compared treated women with untreated subfertile women, no excess risk was noted in association with hyperstimulation medications. Two cohorts noted an increased risk of twofold to fivefold when treated women were compared with the general population. An increased risk of borderline ovarian tumors was noted in three case-control studies and two cohort studies. Overall, the authors concluded there was no convincing evidence that an increased risk of invasive ovarian tumors was associated with fertility drug treatments.[48]

After the Cochrane review, a follow-up study of an infertility cohort [49] was published. A retrospective cohort of 9,825 women enrolled between 1965 and 1988 was followed through 2010. Ovarian cancer occurred in 85 women. Overall, there was no association between ovarian cancer and clomiphene citrate (RR, 1.34; 95% CI, 0.86–2.07) or gonadotropins (RR, 1.00; 95% CI, 0.48–2.08). Among the subgroup of women who remained nulligravid after treatment, an increased risk of ovarian cancer was associated with clomiphene citrate (RR, 3.63; 95% CI, 1.36–9.72); no increased risk was observed among women who successfully conceived after being treated, compared with women who were not treated.

References

1. American Cancer Society: Cancer Facts and Figures 2019. Atlanta, Ga: American Cancer Society, 2019. Available online. Last accessed January 23, 2019.
2. Howlader N, Noone AM, Krapcho M, et al., eds.: SEER Cancer Statistics Review, 1975-2012. Bethesda, Md: National Cancer Institute, 2015. Also available online. Last accessed January 31, 2019.
3. Howlader N, Noone AM, Krapcho M, et al., eds.: SEER Cancer Statistics Review (CSR) 1975-2014. Bethesda, Md: National Cancer Institute. Also available online. Last accessed February 8, 2019.
4. Kurman RJ, Shih IeM: The Dualistic Model of Ovarian Carcinogenesis: Revisited, Revised, and Expanded. Am J Pathol 186 (4): 733-47, 2016. [PUBMED Abstract]
5. Cancer Genome Atlas Research Network: Integrated genomic analyses of ovarian carcinoma. Nature 474 (7353): 609-15, 2011. [PUBMED Abstract]
6. Beral V, Gaitskell K, Hermon C, et al.: Menopausal hormone use and ovarian cancer risk: individual participant meta-analysis of 52 epidemiological studies. Lancet 385 (9980): 1835-42, 2015. [PUBMED Abstract]
7. Lacey JV Jr, Brinton LA, Leitzmann MF, et al.: Menopausal hormone therapy and ovarian cancer risk in the National Institutes of Health-AARP Diet and Health Study Cohort. J Natl Cancer Inst 98 (19): 1397-405, 2006. [PUBMED Abstract]
8. Trabert B, Wentzensen N, Yang HP, et al.: Ovarian cancer and menopausal hormone therapy in the NIH-AARP diet and health study. Br J Cancer 107 (7): 1181-7, 2012. [PUBMED Abstract]
9. Collaborative Group on Epidemiological Studies of Ovarian Cancer: Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. PLoS Med 9 (4): e1001200, 2012. [PUBMED Abstract]
10. Calle EE, Rodriguez C, Walker-Thurmond K, et al.: Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 348 (17): 1625-38, 2003. [PUBMED Abstract]
11. Aune D, Navarro Rosenblatt DA, Chan DS, et al.: Anthropometric factors and ovarian cancer risk: a systematic review and nonlinear dose-response meta-analysis of prospective studies. Int J Cancer 136 (8): 1888-98, 2015. [PUBMED Abstract]
12. Collaborative Group on Epidemiological Studies of Ovarian Cancer, Beral V, Doll R, et al.: Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23,257 women with ovarian cancer and 87,303 controls. Lancet 371 (9609): 303-14, 2008. [PUBMED Abstract]
13. Havrilesky LJ, Moorman PG, Lowery WJ, et al.: Oral contraceptive pills as primary prevention for ovarian cancer: a systematic review and meta-analysis. Obstet Gynecol 122 (1): 139-47, 2013. [PUBMED Abstract]
14. Depot-medroxyprogesterone acetate (DMPA) and risk of epithelial ovarian cancer. The WHO Collaborative Study of Neoplasia and Steroid Contraceptives. Int J Cancer 49 (2): 191-5, 1991. [PUBMED Abstract]
15. Wilailak S, Vipupinyo C, Suraseranivong V, et al.: Depot medroxyprogesterone acetate and epithelial ovarian cancer: a multicentre case-control study. BJOG 119 (6): 672-7, 2012. [PUBMED Abstract]
16. Cibula D, Widschwendter M, Májek O, et al.: Tubal ligation and the risk of ovarian cancer: review and meta-analysis. Hum Reprod Update 17 (1): 55-67, 2011 Jan-Feb. [PUBMED Abstract]
17. Ness RB, Dodge RC, Edwards RP, et al.: Contraception methods, beyond oral contraceptives and tubal ligation, and risk of ovarian cancer. Ann Epidemiol 21 (3): 188-96, 2011. [PUBMED Abstract]
18. Sieh W, Salvador S, McGuire V, et al.: Tubal ligation and risk of ovarian cancer subtypes: a pooled analysis of case-control studies. Int J Epidemiol 42 (2): 579-89, 2013. [PUBMED Abstract]

19. Wentzensen N, Poole EM, Trabert B, et al.: Ovarian Cancer Risk Factors by Histologic Subtype: An Analysis From the Ovarian Cancer Cohort Consortium. J Clin Oncol 34 (24): 2888-98, 2016. [PUBMED Abstract]

20. Luan NN, Wu QJ, Gong TT, et al.: Breastfeeding and ovarian cancer risk: a meta-analysis of epidemiologic studies. Am J Clin Nutr 98 (4): 1020-31, 2013. [PUBMED Abstract]

21. Li DP, Du C, Zhang ZM, et al.: Breastfeeding and ovarian cancer risk: a systematic review and meta-analysis of 40 epidemiological studies. Asian Pac J Cancer Prev 15 (12): 4829-37, 2014. [PUBMED Abstract]

22. Feng LP, Chen HL, Shen MY: Breastfeeding and the risk of ovarian cancer: a meta-analysis. J Midwifery Womens Health 59 (4): 428-37, 2014 Jul-Aug. [PUBMED Abstract]

23. Hanley GE, McAlpine JN, Kwon JS, et al.: Opportunistic salpingectomy for ovarian cancer prevention. Gynecol Oncol Res Pract 2: 5, 2015. [PUBMED Abstract]

24. Daly MB, Dresher CW, Yates MS, et al.: Salpingectomy as a means to reduce ovarian cancer risk. Cancer Prev Res (Phila) 8 (5): 342-8, 2015. [PUBMED Abstract]

25. Duan L, Xu X, Koebnick C, et al.: Bilateral oophorectomy is not associated with increased mortality: the California Teachers Study. Fertil Steril 97 (1): 111-7, 2012. [PUBMED Abstract]

26. Rocca WA, Grossardt BR, de Andrade M, et al.: Survival patterns after oophorectomy in premenopausal women: a population-based cohort study. Lancet Oncol 7 (10): 821-8, 2006. [PUBMED Abstract]

27. McCarthy AM, Menke A, Ouyang P, et al.: Bilateral oophorectomy, body mass index, and mortality in U.S. women aged 40 years and older. Cancer Prev Res (Phila) 5 (6): 847-54, 2012. [PUBMED Abstract]

28. Rivera CM, Grossardt BR, Rhodes DJ, et al.: Increased cardiovascular mortality after early bilateral oophorectomy. Menopause 16 (1): 15-23, 2009 Jan-Feb. [PUBMED Abstract]

29. Parker WH, Feskanich D, Broder MS, et al.: Long-term mortality associated with oophorectomy compared with ovarian conservation in the nurses' health study. Obstet Gynecol 121 (4): 709-16, 2013. [PUBMED Abstract]

30. Jacoby VL, Grady D, Wactawski-Wende J, et al.: Oophorectomy vs ovarian conservation with hysterectomy: cardiovascular disease, hip fracture, and cancer in the Women's Health Initiative Observational Study. Arch Intern Med 171 (8): 760-8, 2011. [PUBMED Abstract]

31. Yoon SH, Kim SN, Shim SH, et al.: Bilateral salpingectomy can reduce the risk of ovarian cancer in the general population: A meta-analysis. Eur J Cancer 55: 38-46, 2016. [PUBMED Abstract]

32. Falconer H, Yin L, Grönberg H, et al.: Ovarian cancer risk after salpingectomy: a nationwide population-based study. J Natl Cancer Inst 107 (2): , 2015. [PUBMED Abstract]

33. Findley AD, Siedhoff MT, Hobbs KA, et al.: Short-term effects of salpingectomy during laparoscopic hysterectomy on ovarian reserve: a pilot randomized controlled trial. Fertil Steril 100 (6): 1704-8, 2013. [PUBMED Abstract]

34. Venturella R, Lico D, Borelli M, et al.: 3 to 5 Years Later: Long-term Effects of Prophylactic Bilateral Salpingectomy on Ovarian Function. J Minim Invasive Gynecol 24 (1): 145-150, 2017. [PUBMED Abstract]

35. Rota M, Pasquali E, Scotti L, et al.: Alcohol drinking and epithelial ovarian cancer risk. a systematic review and meta-analysis. Gynecol Oncol 125 (3): 758-63, 2012. [PUBMED Abstract]

36. Chandran U, Bandera EV, Williams-King MG, et al.: Healthy eating index and ovarian cancer risk. Cancer Causes Control 22 (4): 563-71, 2011. [PUBMED Abstract]

37. Crane TE, Khulpateea BR, Alberts DS, et al.: Dietary intake and ovarian cancer risk: a systematic review. Cancer Epidemiol Biomarkers Prev 23 (2): 255-73, 2014. [PUBMED Abstract]

38. Baandrup L, Faber MT, Christensen J, et al.: Nonsteroidal anti-inflammatory drugs and risk of ovarian cancer: systematic review and meta-analysis of observational studies. Acta Obstet Gynecol Scand 92 (3): 245-55, 2013. [PUBMED Abstract]

39. Murphy MA, Trabert B, Yang HP, et al.: Non-steroidal anti-inflammatory drug use and ovarian cancer risk: findings from the NIH-AARP Diet and Health Study and systematic review. Cancer Causes Control 23 (11): 1839-52, 2012. [PUBMED Abstract]

40. Lo-Ciganic WH, Zgibor JC, Bunker CH, et al.: Aspirin, nonaspirin nonsteroidal anti-inflammatory drugs, or acetaminophen and risk of ovarian cancer. Epidemiology 23 (2): 311-9, 2012. [PUBMED Abstract]

41. Barnard ME, Poole EM, Curhan GC, et al.: Association of Analgesic Use With Risk of Ovarian Cancer in the Nurses' Health Studies. JAMA Oncol 4 (12): 1675-1682, 2018. [PUBMED Abstract]

42. Huncharek M, Geschwind JF, Kupelnick B: Perineal application of cosmetic talc and risk of invasive epithelial ovarian cancer: a meta-analysis of 11,933 subjects from sixteen observational studies. Anticancer Res 23 (2C): 1955-60, 2003 Mar-Apr. [PUBMED Abstract]

43. Terry KL, Karageorgi S, Shvetsov YB, et al.: Genital powder use and risk of ovarian cancer: a pooled analysis of 8,525 cases and 9,859 controls. Cancer Prev Res (Phila) 6 (8): 811-21, 2013. [PUBMED Abstract]

44. Schildkraut JM, Abbott SE, Alberg AJ, et al.: Association between Body Powder Use and Ovarian Cancer: The African American Cancer Epidemiology Study (AACES). Cancer Epidemiol Biomarkers Prev 25 (10): 1411-1417, 2016. [PUBMED Abstract]

45. Gertig DM, Hunter DJ, Cramer DW, et al.: Prospective study of talc use and ovarian cancer. J Natl Cancer Inst 92 (3): 249-52, 2000. [PUBMED Abstract]

46. Houghton SC, Reeves KW, Hankinson SE, et al.: Perineal powder use and risk of ovarian cancer. J Natl Cancer Inst 106 (9): , 2014. [PUBMED Abstract]

47. Siristatidis C, Sergentanis TN, Kanavidis P, et al.: Controlled ovarian hyperstimulation for IVF: impact on ovarian, endometrial and cervical cancer--a systematic review and meta-analysis. Hum Reprod Update 19 (2): 105-23, 2013 Mar-Apr. [PUBMED Abstract]

48. Rizzuto I, Behrens RF, Smith LA: Risk of ovarian cancer in women treated with ovarian stimulating drugs for infertility. Cochrane Database Syst Rev 8: CD008215, 2013. [PUBMED Abstract]

49. Trabert B, Lamb EJ, Scoccia B, et al.: Ovulation-inducing drugs and ovarian cancer risk: results from an extended follow-up of a large United States infertility cohort. Fertil Steril 100 (6): 1660-6, 2013. [PUBMED Abstract]

Changes to This Summary (03/01/2019)

The PDQ cancer information summaries are reviewed regularly and updated as new information becomes available. This section describes the latest changes made to this summary as of the date above.

Description of the Evidence

Updated statistics with estimated new cases and deaths for 2019 (cited American Cancer Society as reference 1).

Added text about a cohort analysis of about 200,000 women in the Nurses' Health Studies, which used detailed data about the intensity and duration of aspirin use over time, that showed a reduced hazard ratio for ovarian cancer of 0.77 for low-dose aspirin use but no reduction for standard-dose aspirin use (cited Barnard et al. as reference 41).

This summary is written and maintained by the PDQ Screening and Prevention Editorial Board, which is editorially independent of NCI. The summary reflects an independent review of the literature and does not represent a policy statement of NCI or NIH. More information about summary policies and the role of the PDQ Editorial Boards in maintaining the PDQ summaries can be found on the About This PDQ Summary and PDQ® - NCI's Comprehensive Cancer Database pages.

About This PDQ Summary

Purpose of This Summary

This PDQ cancer information summary for health professionals provides comprehensive, peer-reviewed, evidence-based information about ovarian, fallopian tube, and primary peritoneal cancer prevention. It is intended as a resource to inform and assist clinicians who care for cancer patients. It does not provide formal guidelines or recommendations for making health care decisions.

Reviewers and Updates

This summary is reviewed regularly and updated as necessary by the PDQ Screening and Prevention Editorial Board, which is editorially independent of the National Cancer Institute (NCI). The summary reflects an independent review of the literature and does not represent a policy statement of NCI or the National Institutes of Health (NIH).

Board members review recently published articles each month to determine whether an article should:

- be discussed at a meeting,
- be cited with text, or
- replace or update an existing article that is already cited.

Changes to the summaries are made through a consensus process in which Board members evaluate the strength of the evidence in the published articles and determine how the article should be included in the summary.

Any comments or questions about the summary content should be submitted to Cancer.gov through the NCI website's Email Us. Do not contact the individual Board Members with questions or comments about the summaries. Board members will not respond to individual inquiries.

Levels of Evidence

Some of the reference citations in this summary are accompanied by a level-of-evidence designation. These designations are intended to help readers assess the strength of the evidence supporting the use of specific interventions or approaches. The PDQ Screening and Prevention Editorial Board uses a formal evidence ranking system in developing its level-of-evidence designations.

Permission to Use This Summary

PDQ is a registered trademark. Although the content of PDQ documents can be used freely as text, it cannot be identified as an NCI PDQ cancer information summary unless it is presented in its entirety and is regularly updated. However, an author would be permitted to write a sentence such as "NCI's PDQ cancer information summary about breast cancer prevention states the risks succinctly: [include excerpt from the summary]."

The preferred citation for this PDQ summary is:

PDQ® Screening and Prevention Editorial Board. PDQ Ovarian, Fallopian Tube, and Primary Peritoneal Cancer Prevention. Bethesda, MD: National Cancer Institute. Updated <MM/DD/YYYY>. Available at: https://www.cancer.gov/types/ovarian/hp/ovarian-prevention-pdq. Accessed <MM/DD/YYYY>. [PMID: 26389359]

Images in this summary are used with permission of the author(s), artist, and/or publisher for use within the PDQ summaries only. Permission to use images outside the context of PDQ information must be obtained from the owner(s) and cannot be granted by the National Cancer Institute. Information about using the illustrations in this summary, along with many other cancer-related images, is available in Visuals Online, a collection of over 2,000 scientific images.

Disclaimer

The information in these summaries should not be used as a basis for insurance reimbursement determinations. More information on insurance coverage is available on Cancer.gov on the Managing Cancer Care page.

Contact Us

More information about contacting us or receiving help with the Cancer.gov website can be found on our Contact Us for Help page. Questions can also be submitted to Cancer.gov through the website's Email Us.

Updated: March 1, 2019

Exhibit 114

time of the female cycle, normal sperm are able to move through the cervix, but sperm with abnormal tails or impaired swimming ability are detained. These latter sperm then die and are reabsorbed or lost from the body. Other sperm enter *cervical crypts* (deep recesses in the cervical wall), where they die or are lost, or they may remain alive as a reservoir of sperm that later may enter the uterus. Fewer than 1 million of the original approximately 200 million sperm make it through the cervix.

Uterine Sperm

Upon leaving the cervix, the sperm travel up the uterus to the uterotubal junction. The uterine fluid is watery but sparse in humans, and the sperm essentially "climb" up the uterine lumen by beating their tails. The swimming rate of sperm (about 3 mm/min), however, cannot account for their traveling a distance of about 15 cm in the 30 min after ejaculation. Also, dead sperm reach the oviduct at about the same time as do live sperm. Thus, sperm tail beating probably is not important during sperm transport through the uterus, so it must be the muscle contraction and movement of cilia in the female reproductive tract that facilitate sperm transport.

During the later follicular phase, a thin layer of myometrium just underneath the uterine endometrium contracts, directing fluid toward the fallopian tubes. If coitus has occurred, these contractions probably increase. Mechanical stimulation of the cervix by the penis during coitus causes release of the hormone oxytocin from a woman's posterior pituitary gland. This hormone quickly travels via the blood to the uterus and increases the force of rhythmic uterine muscle contractions. These contractions act as waves to help the sperm move to the uterotubal junction. Prostaglandins in the seminal fluid may also cause uterine muscles to contract, but this is unlikely as very little seminal fluid enters the uterus through the cervix. The main function of the prostaglandins in seminal fluid is probably to contract the muscles of the vasa deferentia, thus aiding sperm passage during ejaculation.

The vagina, cervix, and uterus respond to the presence of sperm with an immunological attack, and those sperm cells that do not move rapidly through these portions of the female tract are destroyed. Thus, the uterotubal junction allows the gradual entrance of sperm into the isthmus of the oviduct. About half of the sperm enter the wrong oviduct, and only a few hundred make it to the general proximity of the waiting egg.

TRANSPORT OF THE SPERM AND OVUM IN THE OVIDUCT

Sperm that survive the journey into the fallopian tube find a less hostile environment than in other portions of the female tract. Attack by leukocytes is much reduced. The passage of sperm slows down as the cells encounter mucus in the lumen of the oviduct and have to navigate the increasingly complex folds of the inner lining of the tube. Evidence suggests that sperm heads make contact with and briefly bind to the epithelial cells of the mucosal lining. This contact somehow preserves and extends the viability of the sperm. By slowing down the progress of the sperm, the oviducts may serve as a reservoir from which sperm may gradually proceed, thus extending the time that sperm are available to fertilize an egg. After ovulation, sperm approach the ovum, and fertilization by a single sperm usually occurs at the point where the isthmus joins the wider oviductal ampulla (ampullary–isthmic junction). Other sperm swim up the ampulla, through the infundibulum, and are lost in the body cavity.

Once ovulation has occurred, the infundibulum (funnel-shaped free end) of the oviduct moves toward the ovary and envelops the ovulated ovum along with fluid derived from the ovulated follicle. Movement of the infundibulum is accomplished by the contraction of muscles in the membrane supporting the oviduct. Cilia are present in the wall of the fimbria (the edge of the infundibulum), and these beat toward the uterus. Thus, when the infundibulum envelops the ovary, the beating of the cilia moves the ovum into the ampulla of the oviduct. Cilia in the ampulla and isthmus of the oviduct also beat in a uterine direction, which sets up a flow of fluid toward the uterus.

The muscles of the oviduct also exhibit waves of muscular contraction after ovulation. These waves travel in the direction of the uterus and, along with the cilia, help

Exhibit 115

The Environment and Disease: Association or Causation?

by Sir Austin Bradford Hill CBE DSC FRCP(hon) FRS
(*Professor Emeritus of Medical Statistics, University of London*)

Amongst the objects of this newly-founded Section of Occupational Medicine are firstly 'to provide a means, not readily afforded elsewhere, whereby physicians and surgeons with a special knowledge of the relationship between sickness and injury and conditions of work may discuss their problems, not only with each other, but also with colleagues in other fields, by holding joint meetings with other Sections of the Society'; and, secondly, 'to make available information about the physical, chemical and psychological hazards of occupation, and in particular about those that are rare or not easily recognized'.

At this first meeting of the Section and before, with however laudable intentions, we set about instructing our colleagues in other fields, it will be proper to consider a problem fundamental to our own. How in the first place do we detect these relationships between sickness, injury and conditions of work? How do we determine what are physical, chemical and psychological hazards of occupation, and in particular those that are rare and not easily recognized?

There are, of course, instances in which we can reasonably answer these questions from the general body of medical knowledge. A particular, and perhaps extreme, physical environment cannot fail to be harmful; a particular chemical is known to be toxic to man and therefore suspect on the factory floor. Sometimes, alternatively, we may be able to consider what *might* a particular environment do to man, and then see whether such consequences are indeed to be found. But more often than not we have no such guidance, no such means of proceeding; more often than not we are dependent upon our observation and enumeration of defined events for which we then seek antecedents. In other words we see that the event B is associated with the environmental feature A, that, to take a specific example, some form of respiratory illness is associated with a dust in the environment. In what circumstances can we pass from this

Meeting January 14 1965

President's Address

observed *association* to a verdict of *causation*? Upon what basis should we proceed to do so?

I have no wish, nor the skill, to embark upon a philosophical discussion of the meaning of 'causation'. The 'cause' of illness may be immediate and direct, it may be remote and indirect underlying the observed association. But with the aims of occupational, and almost synonymously preventive, medicine in mind the decisive question is whether the frequency of the undesirable event B will be influenced by a change in the environmental feature A. *How* such a change exerts that influence may call for a great deal of research. However, before deducing 'causation' and taking action we shall not invariably have to sit around awaiting the results of that research. The whole chain may have to be unravelled or a few links may suffice. It will depend upon circumstances.

Disregarding then any such problem in semantics we have this situation. Our observations reveal an association between two variables, perfectly clear-cut and beyond what we would care to attribute to the play of chance. What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?

(1) *Strength*. First upon my list I would put the strength of the association. To take a very old example, by comparing the occupations of patients with scrotal cancer with the occupations of patients presenting with other diseases, Percival Pott could reach a correct conclusion because of the *enormous* increase of scrotal cancer in the chimney sweeps. 'Even as late as the second decade of the twentieth century', writes Richard Doll (1964), 'the mortality of chimney sweeps from scrotal cancer was some 200 times that of workers who were not specially exposed to tar or mineral oils and in the eighteenth century the relative difference is likely to have been much greater.'

To take a more modern and more general example upon which I have now reflected for over fifteen years, prospective inquiries into smoking have shown that the death rate from cancer of the lung in cigarette smokers is nine to ten times the rate in non-smokers and the rate in heavy cigarette smokers is twenty to thirty times

as great. On the other hand the death rate from coronary thrombosis in smokers is no more than twice, possibly less, the death rate in non-smokers. Though there is good evidence to support causation it is surely much easier in this case to think of some features of life that may go hand-in-hand with smoking – features that might conceivably be the real underlying cause or, at the least, an important contributor, whether it be lack of exercise, nature of diet or other factors. But to explain the pronounced excess in cancer of the lung in any other environmental terms requires some feature of life so intimately linked with cigarette smoking and with the amount of smoking that such a feature should be easily detectable. If we cannot detect it or reasonably infer a specific one, then in such circumstances I think we are reasonably entitled to reject the vague contention of the armchair critic 'you can't prove it, there *may* be such a feature'.

Certainly in this situation I would reject the argument sometimes advanced that what matters is the absolute difference between the death rates of our various groups and not the ratio of one to other. That depends upon what we want to know. If we want to know how many extra deaths from cancer of the lung will take place through smoking (i.e. presuming causation), then obviously we must use the absolute differences between the death rates – 0·07 per 1,000 per year in non-smoking doctors, 0·57 in those smoking 1–14 cigarettes daily, 1·39 for 15–24 cigarettes daily and 2·27 for 25 or more daily. But it does not follow here, or in more specifically occupational problems, that this best measure of the effect upon mortality is also the best measure in relation to ætiology. In this respect the ratios of 8, 20 and 32 to 1 are far more informative. It does not, of course, follow that the differences revealed by ratios are of any practical importance. Maybe they are, maybe they are not; but that is another point altogether.

We may recall John Snow's classic analysis of the opening weeks of the cholera epidemic of 1854 (Snow 1855). The death rate that he recorded in the customers supplied with the grossly polluted water of the Southwark and Vauxhall Company was in truth quite low – 71 deaths in each 10,000 houses. What stands out vividly is the fact that the small rate is 14 times the figure of 5 deaths per 10,000 houses supplied with the sewage-free water of the rival Lambeth Company.

In thus putting emphasis upon the strength of an association we must, nevertheless, look at the obverse of the coin. We must not be too ready to dismiss a cause-and-effect hypothesis merely on the grounds that the observed association appears to be slight. There are many occasions in medicine when this is in truth so. Relatively few persons harbouring the meningococcus fall sick of meningococcal meningitis. Relatively few persons occupationally exposed to rat's urine contract Weil's disease.

(2) *Consistency:* Next on my list of features to be specially considered I would place the *consistency* of the observed association. Has it been repeatedly observed by different persons, in different places, circumstances and times?

This requirement may be of special importance for those rare hazards singled out in the Section's terms of reference. With many alert minds at work in industry today many an environmental association may be thrown up. Some of them on the customary tests of statistical significance will appear to be unlikely to be due to chance. Nevertheless whether chance is the explanation or whether a true hazard has been revealed may sometimes be answered only by a repetition of the circumstances and the observations.

Returning to my more general example, the Advisory Committee to the Surgeon-General of the United States Public Health Service found the association of smoking with cancer of the lung in 29 retrospective and 7 prospective inquiries (US Department of Health, Education & Welfare 1964). The lesson here is that broadly the same answer has been reached in quite a wide variety of situations and techniques. In other words we can justifiably infer that the association is not due to some constant error or fallacy that permeates every inquiry. And we have indeed to be on our guard against that.

Take, for instance, an example given by Heady (1958). Patients admitted to hospital for operation for peptic ulcer are questioned about recent domestic anxieties or crises that may have precipitated the acute illness. As controls, patients admitted for operation for a simple hernia are similarly quizzed. But, as Heady points out, the two groups may not be *in pari materia*. If your wife ran off with the lodger last week you still have to take your perforated ulcer to hospital without delay. But with a hernia you might prefer to stay at home for a while – to mourn (or celebrate) the event. No number of exact repetitions would remove or necessarily reveal that fallacy.

We have, therefore, the somewhat paradoxical position that the different results of a different inquiry certainly cannot be held to refute the

original evidence; yet the same results from precisely the same form of inquiry will not invariably greatly strengthen the original evidence. I would myself put a good deal of weight upon similar results reached in quite different ways, e.g. prospectively and retrospectively.

Once again looking at the obverse of the coin there will be occasions when repetition is absent or impossible and yet we should not hesitate to draw conclusions. The experience of the nickel refiners of South Wales is an outstanding example. I quote from the Alfred Watson Memorial Lecture that I gave in 1962 to the Institute of Actuaries:

'The population at risk, workers and pensioners, numbered about one thousand. During the ten years 1929 to 1938, sixteen of them had died from cancer of the lung, eleven of them had died from cancer of the nasal sinuses. At the age specific death rates of England and Wales at that time, one might have anticipated one death from cancer of the lung (to compare with the 16), and a fraction of a death from cancer of the nose (to compare with the 11). In all other bodily sites cancer had appeared on the death certificate 11 times and one would have expected it to do so 10–11 times. There had been 67 deaths from all other causes of mortality and over the ten years' period 72 would have been expected at the national death rates. Finally division of the population at risk in relation to their jobs showed that the excess of cancer of the lung and nose had fallen wholly upon the workers employed in the chemical processes.

'More recently my colleague, Dr Richard Doll, has brought this story a stage further. In the nine years 1948 to 1956 there had been, he found, 48 deaths from cancer of the lung and 13 deaths from cancer of the nose. He assessed the numbers expected at normal rates of mortality as, respectively 10 and 0·1.

'In 1923, long before any special hazard had been recognized, certain changes in the refinery took place. No case of cancer of the nose has been observed in any man who first entered the works after that year, and in these men there has been no excess of cancer of the lung. In other words, the excess in both sites is uniquely a feature in men who entered the refinery in, roughly, the first 23 years of the present century.

'No causal agent of these neoplasms has been identified. Until recently no animal experimentation had given any clue or any support to this wholly statistical evidence. Yet I wonder if any of us would hesitate to accept it as proof of a grave industrial hazard?' (Hill 1962).

In relation to my present discussion I know of no parallel investigation. We have (or certainly had) to make up our minds on a unique event; and there is no difficulty in doing so.

(3) *Specificity:* One reason, needless to say, is the specificity of the association, the third characteristic which invariably we must consider. If, as here, the association is limited to specific workers and to particular sites and types of disease and there is no association between the work and other modes of dying, then clearly that is a strong argument in favour of causation.

We must not, however, over-emphasize the importance of the characteristic. Even in my present example there is a cause and effect relationship with two different sites of cancer – the lung and the nose. Milk as a carrier of infection and, in that sense, the cause of disease can produce such a disparate galaxy as scarlet fever, diphtheria, tuberculosis, undulant fever, sore throat, dysentery and typhoid fever. Before the discovery of the underlying factor, the bacterial origin of disease, harm would have been done by pushing too firmly the need for specificity as a necessary feature before convicting the dairy.

Coming to modern times the prospective investigations of smoking and cancer of the lung have been criticized for not showing specificity – in other words the death rate of smokers is higher than the death rate of non-smokers from many causes of death (though in fact the results of Doll & Hill, 1964, do not show that). But here surely one must return to my first characteristic, the strength of the association. If other causes of death are raised 10, 20 or even 50% in smokers whereas cancer of the lung is raised 900–1,000% we have specificity – a specificity in the magnitude of the association.

We must also keep in mind that diseases may have more than one cause. It has always been possible to acquire a cancer of the scrotum without sweeping chimneys or taking to mule-spinning in Lancashire. One-to-one relationships are not frequent. Indeed I believe that multi-causation is generally more likely than single causation though possibly if we knew all the answers we might get back to a single factor.

In short, if specificity exists we may be able to draw conclusions without hesitation; if it is not apparent, we are not thereby necessarily left sitting irresolutely on the fence.

(4) *Temporality:* My fourth characteristic is the temporal relationship of the association – which is the cart and which the horse? This is a question which might be particularly relevant with diseases of slow development. Does a particular diet lead to disease or do the early stages of the disease lead to those peculiar dietetic habits? Does a

particular occupation or occupational environment promote infection by the tubercle bacillus or are the men and women who select that kind of work more liable to contract tuberculosis whatever the environment – or, indeed, have they already contracted it? This temporal problem may not arise often but it certainly needs to be remembered, particularly with selective factors at work in industry.

(5) *Biological gradient:* Fifthly, if the association is one which can reveal a biological gradient, or dose-response curve, then we should look most carefully for such evidence. For instance, the fact that the death rate from cancer of the lung rises linearly with the number of cigarettes smoked daily, adds a very great deal to the simpler evidence that cigarette smokers have a higher death rate than non-smokers. That comparison would be weakened, though not necessarily destroyed, if it depended upon, say, a much heavier death rate in light smokers and a lower rate in heavier smokers. We should then need to envisage some much more complex relationship to satisfy the cause-and-effect hypothesis. The clear dose-response curve admits of a simple explanation and obviously puts the case in a clearer light.

The same would clearly be true of an alleged dust hazard in industry. The dustier the environment the greater the incidence of disease we would expect to see. Often the difficulty is to secure some satisfactory quantitative measure of the environment which will permit us to explore this dose-response. But we should invariably seek it.

(6) *Plausibility:* It will be helpful if the causation we suspect is biologically plausible. But this is a feature I am convinced we cannot demand. What is biologically plausible depends upon the biological knowledge of the day.

To quote again from my Alfred Watson Memorial Lecture (Hill 1962), there was

'. . . no biological knowledge to support (or to refute) Pott's observation in the 18th century of the excess of cancer in chimney sweeps. It was lack of biological knowledge in the 19th that led a prize essayist writing on the value and the fallacy of statistics to conclude, amongst other "absurd" associations, that "it could be no more ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he there contracted, to the vermin with which bodies of the sick might be infected". And coming to nearer times, in the 20th century there was no biological knowledge to support the evidence against rubella.'

In short, the association we observe may be one new to science or medicine and we must not dismiss it too light-heartedly as just too odd. As Sherlock Holmes advised Dr Watson, 'when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth.'

(7) *Coherence:* On the other hand the cause-and-effect interpretation of our data should not seriously conflict with the generally known facts of the natural history and biology of the disease – in the expression of the Advisory Committee to the Surgeon-General it should have coherence.

Thus in the discussion of lung cancer the Committee finds its association with cigarette smoking coherent with the temporal rise that has taken place in the two variables over the last generation and with the sex difference in mortality – features that might well apply in an occupational problem. The known urban/rural ratio of lung cancer mortality does not detract from coherence, nor the restriction of the effect to the lung.

Personally, I regard as greatly contributing to coherence the histopathological evidence from the bronchial epithelium of smokers and the isolation from cigarette smoke of factors carcinogenic for the skin of laboratory animals. Nevertheless, while such laboratory evidence can enormously strengthen the hypothesis and, indeed, may determine the actual causative agent, the lack of such evidence cannot nullify the epidemiological observations in man. Arsenic can undoubtedly cause cancer of the skin in man but it has never been possible to demonstrate such an effect on any other animal. In a wider field John Snow's epidemiological observations on the conveyance of cholera by the water from the Broad Street pump would have been put almost beyond dispute if Robert Koch had been then around to isolate the vibrio from the baby's nappies, the well itself and the gentleman in delicate health from Brighton. Yet the fact that Koch's work was to be awaited another thirty years did not really weaken the epidemiological case though it made it more difficult to establish against the criticisms of the day – both just and unjust.

(8) *Experiment:* Occasionally it is possible to appeal to experimental, or semi-experimental, evidence. For example, because of an observed association some preventive action is taken. Does it in fact prevent? The dust in the workshop is reduced, lubricating oils are changed, persons stop smoking cigarettes. Is the frequency of the associated events affected? Here the strongest

support for the causation hypothesis may be revealed.

(9) *Analogy:* In some circumstances it would be fair to judge by analogy. With the effects of thalidomide and rubella before us we would surely be ready to accept slighter but similar evidence with another drug or another viral disease in pregnancy.

Here then are nine different viewpoints from all of which we should study association before we cry causation. What I do not believe – and this has been suggested – is that we can usefully lay down some hard-and-fast rules of evidence that *must* be obeyed before we accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?

Tests of Significance

No formal tests of significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us in the likely magnitude of those effects. Beyond that they contribute nothing to the 'proof' of our hypothesis.

Nearly forty years ago, amongst the studies of occupational health that I made for the Industrial Health Research Board of the Medical Research Council was one that concerned the workers in the cotton-spinning mills of Lancashire (Hill 1930). The question that I had to answer, by the use of the National Health Insurance records of that time, was this: Do the workers in the card-room of the spinning mill, who tend the machines that clean the raw cotton, have a sickness experience in any way different from that of other operatives in the same mills who are relatively unexposed to the dust and fibre that were features of the cardroom? The answer was an unqualified 'Yes'. From age 30 to age 60 the cardroom workers suffered over three times as much from respiratory causes of illness whereas from non-respiratory causes their experience was not different from that of the other workers. This pronounced difference with the respiratory causes was derived not from abnormally long periods of sickness but rather from an excessive number of repeated absences from work of the cardroom workers.

All this has rightly passed into the limbo of forgotten things. What interests me today is this: My results were set out for men and women separately and for half a dozen age groups in 36 tables. So there were plenty of sums. Yet I cannot find that anywhere I thought it necessary to use a test of significance. The evidence was so clear-cut, the differences between the groups were mainly so large, the contrast between respiratory and non-respiratory causes of illness so specific, that no formal tests could really contribute anything of value to the argument. So why use them?

Would we think or act that way today? I rather doubt it. Between the two world wars there was a strong case for emphasizing to the clinician and other research workers the importance of not overlooking the effects of the play of chance upon their data. Perhaps too often generalities were based upon two men and a laboratory dog while the treatment of choice was deduced from a difference between two bedfuls of patients and might easily have no true meaning. It was therefore a useful corrective for statisticians to stress, and to teach the need for, tests of significance merely to serve as guides to caution before drawing a conclusion, before inflating the particular to the general.

I wonder whether the pendulum has not swung too far – not only with the attentive pupils but even with the statisticians themselves. To decline to draw conclusions without standard errors can surely be just as silly? Fortunately I believe we have not yet gone so far as our friends in the USA where, I am told, some editors of journals will return an article because tests of significance have not been applied. Yet there are innumerable situations in which they are totally unnecessary – because the difference is grotesquely obvious, because it is negligible, or because, whether it be formally significant or not, it is too small to be of any practical importance. What is worse the glitter of the t table diverts attention from the inadequacies of the fare. Only a tithe, and an unknown tithe, of the factory personnel volunteer for some procedure or interview, 20% of patients treated in some particular way are lost to sight, 30% of a randomly-drawn sample are never contacted. The sample may, indeed, be akin to that of the man who, according to Swift, 'had a mind to sell his house and carried a piece of brick in his pocket, which he showed as a pattern to encourage purchasers'. The writer, the editor and the reader are unmoved. The magic formulæ are there.

Of course I exaggerate. Yet too often I suspect we waste a deal of time, we grasp the shadow and

lose the substance, we weaken our capacity to interpret data and to take reasonable decisions whatever the value of P. And far too often we deduce 'no difference' from 'no significant difference'. Like fire, the χ^2 test is an excellent servant and a bad master.

The Case for Action

Finally, in passing from association to causation I believe in 'real life' we shall have to consider what flows from that decision. On scientific grounds we should do no such thing. The evidence is there to be judged on its merits and the judgment (in that sense) should be utterly independent of what hangs upon it – or who hangs because of it. But in another and more practical sense we may surely ask what is involved in our decision. In occupational medicine our object is usually to take action. If this be operative cause and that be deleterious effect, then we shall wish to intervene to abolish or reduce death or disease.

While that is a commendable ambition it almost inevitably leads us to introduce differential standards before we convict. Thus on relatively slight evidence we might decide to restrict the use of a drug for early-morning sickness in pregnant women. If we are wrong in deducing causation from association no great harm will be done. The good lady and the pharmaceutical industry will doubtless survive.

On fair evidence we might take action on what appears to be an occupational hazard, e.g. we might change from a probably carcinogenic oil to a non-carcinogenic oil in a limited environment and without too much injustice if we are wrong. But we should need very strong evidence before we made people burn a fuel in their homes that they do not like or stop smoking the cigarettes and eating the fats and sugar that they do like. In asking for very strong evidence I would, however, repeat emphatically that this does not imply crossing every 't', and swords with every critic, before we act.

All scientific work is incomplete – whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.

Who knows, asked Robert Browning, but the world may end tonight? True, but on available evidence most of us make ready to commute on the 8.30 next day.

REFERENCES

Doll R (1964) In: Medical Surveys and Clinical Trials. Ed. L J Witts. 2nd ed. London; p 333

Doll R & Hill A B (1964) *Brit. med. J.* i, 1399, 1460

Heady J A (1958) *Med. World, Lond.* 89, 305

Hill A B
(1930) Sickness amongst Operatives in Lancashire Spinning Mills. Industrial Health Research Board Report No. 59. HMSO, London
(1962) *J. Inst. Actu.* 88, 178

Snow J (1855) On the Mode of Communication of Cholera. 2nd ed. London (Reprinted 1936, New York)

US Department of Health, Education & Welfare (1964) Smoking and Health. Public Health Service Publication No. 1103. Washington

Exhibit 116

THIRD EDITION

MODERN EPIDEMIOLOGY

Kenneth J. Rothman
Vice President, Epidemiology Research
RTI Health Solutions
Professor of Epidemiology and Medicine
Boston University
Boston, Massachusetts

Sander Greenland
Professor of Epidemiology and Statistics
University of California
Los Angeles, California

Timothy L. Lash
Associate Professor of Epidemiology and Medicine
Boston University
Boston, Massachusetts

Wolters Kluwer Health | Lippincott Williams & Wilkins
Philadelphia • Baltimore • New York • London
Buenos Aires • Hong Kong • Sydney • Tokyo

Acquisitions Editor: Sonya Seigafuse
Developmental Editor: Louise Bierig
Project Manager: Kevin Johnson
Senior Manufacturing Manager: Ben Rivera
Marketing Manager: Kimberly Schonberger
Art Director: Risa Clow
Compositor: Aptara, Inc.

530 Walnut Street
Philadelphia, PA 19106 USA
LWW.com

Printed in China

Library of Congress Cataloging-in-Publication Data
Rothman, Kenneth J.
Modern epidemiology / Kenneth J. Rothman, Sander Greenland, and Timothy L. Lash. – 3rd ed.
p. ; cm.
2nd ed. edited by Kenneth J. Rothman and Sander Greenland.
Includes bibliographical references and index.
ISBN-13: 978-0-7817-5564-1
ISBN-10: 0-7817-5564-6
1. Epidemiology–Statistical methods. 2. Epidemiology--Research–Methodology. I. Greenland, Sander, 1951- II. Lash, Timothy L. III. Title.
[DNLM: 1. Epidemiology. 2. Epidemiologic Methods. WA 105 R846m 2008]
RA652.2.M3R67 2008
614.4–dc22
2007036316

10 9 8 7 6 5 4

RRS1105

Contents

SECTION III
Data Analysis

SECTION IV
Special Topics

Preface and Acknowledgments

This third edition of *Modern Epidemiology* arrives more than 20 years after the first edition, which was a much smaller single-authored volume that outlined the concepts and methods of a rapidly growing discipline. The second edition, published 12 years later, was a major transition, as the book grew along with the field. It saw the addition of a second author and an expansion of topics contributed by invited experts in a range of subdisciplines. Now, with the help of a third author, this new edition encompasses a comprehensive revision of the content and the introduction of new topics that 21st century epidemiologists will find essential.

This edition retains the basic organization of the second edition, with the book divided into four parts. Part I (Basic Concepts) now comprises five chapters rather than four, with the relocation of Chapter 5, "Concepts of Interaction," which was Chapter 18 in the second edition. The topic of interaction rightly belongs with Basic Concepts, although a reader aiming to accrue a working understanding of epidemiologic principles could defer reading it until after Part II, "Study Design and Conduct." We have added a new chapter on causal diagrams, which we debated putting into Part I, as it does involve basic issues in the conceptualization of relations between study variables. On the other hand, this material invokes concepts that seemed more closely linked to data analysis, and assumes knowledge of study design, so we have placed it at the beginning of Part III, "Data Analysis." Those with basic epidemiologic background could read Chapter 12 in tandem with Chapters 2 and 4 to get a thorough grounding in the concepts surrounding causal and non-causal relations among variables. Another important addition is a chapter in Part III titled, "Introduction to Bayesian Statistics," which we hope will stimulate epidemiologists to consider and apply Bayesian methods to epidemiologic settings. The former chapter on sensitivity analysis, now entitled "Bias Analysis," has been substantially revised and expanded to include probabilistic methods that have entered epidemiology from the fields of risk and policy analysis. The rigid application of frequentist statistical interpretations to data has plagued biomedical research (and many other sciences as well). We hope that the new chapters in Part III will assist in liberating epidemiologists from the shackles of frequentist statistics, and open them to more flexible, realistic, and deeper approaches to analysis and inference.

As before, Part IV comprises additional topics that are more specialized than those considered in the first three parts of the book. Although field methods still have wide application in epidemiologic research, there has been a surge in epidemiologic research based on existing data sources, such as registries and medical claims data. Thus, we have moved the chapter on field methods from Part II into Part IV, and we have added a chapter entitled, "Using Secondary Data." Another addition is a chapter on social epidemiology, and coverage on molecular epidemiology has been added to the chapter on genetic epidemiology. Many of these chapters may be of interest mainly to those who are focused on a particular area, such as reproductive epidemiology or infectious disease epidemiology, which have distinctive methodologic concerns, although the issues raised are well worth considering for any epidemiologist who wishes to master the field. Topics such as ecologic studies and meta-analysis retain a broad interest that cuts across subject matter subdisciplines. Screening had its own chapter in the second edition; its content has been incorporated into the revised chapter on clinical epidemiology.

The scope of epidemiology has become too great for a single text to cover it all in depth. In this book, we hope to acquaint those who wish to understand the concepts and methods of epidemiology with the issues that are central to the discipline, and to point the way to key references for further study. Although previous editions of the book have been used as a course text in many epidemiology

teaching programs, it is not written as a text for a specific course, nor does it contain exercises or review questions as many course texts do. Some readers may find it most valuable as a reference or supplementary-reading book for use alongside shorter textbooks such as Kelsey et al. (1996), Szklo and Nieto (2000), Savitz (2001), Koepsell and Weiss (2003), or Checkoway et al. (2004). Nonetheless, there are subsets of chapters that could form the textbook material for epidemiologic methods courses. For example, a course in epidemiologic theory and methods could be based on Chapters 1 through 12, with a more abbreviated course based on Chapters 1 through 4 and 6 through 11. A short course on the foundations of epidemiologic theory could be based on Chapters 1 through 5 and Chapter 12. Presuming a background in basic epidemiology, an introduction to epidemiologic data analysis could use Chapters 9, 10, and 12 through 19, while a more advanced course detailing causal and regression analysis could be based on Chapters 2 through 5, 9, 10, and 12 through 21. Many of the other chapters would also fit into such suggested chapter collections, depending on the program and the curriculum.

Many topics are discussed in various sections of the text because they pertain to more than one aspect of the science. To facilitate access to all relevant sections of the book that relate to a given topic, we have indexed the text thoroughly. We thus recommend that the index be consulted by those wishing to read our complete discussion of specific topics.

We hope that this new edition provides a resource for teachers, students, and practitioners of epidemiology. We have attempted to be as accurate as possible, but we recognize that any work of this scope will contain mistakes and omissions. We are grateful to readers of earlier editions who have brought such items to our attention. We intend to continue our past practice of posting such corrections on an internet page, as well as incorporating such corrections into subsequent printings. Please consult <http://www.lww.com/ModernEpidemiology> to find the latest information on errata.

We are also grateful to many colleagues who have reviewed sections of the current text and provided useful feedback. Although we cannot mention everyone who helped in that regard, we give special thanks to Onyebuchi Arah, Matthew Fox, Jamie Gradus, Jennifer Hill, Katherine Hoggatt, Marshal Joffe, Ari Lipsky, James Robins, Federico Soldani, Henrik Toft Sørensen, Soe Soe Thwin and Tyler VanderWeele. An earlier version of Chapter 18 appeared in the *International Journal of Epidemiology* (2006;35:765–778), reproduced with permission of Oxford University Press. Finally, we thank Mary Anne Armstrong, Alan Dyer, Gary Friedman, Ulrik Gerdes, Paul Sorlie, and Katsuhiko Yano for providing unpublished information used in the examples of Chapter 33.

Kenneth J. Rothman
Sander Greenland
Timothy L. Lash

ıor does it contain exercises or
it most valuable as a reference
s such as Kelsey et al. (1996),
), or Checkoway et al. (2004).
ook material for epidemiologic
ınd methods could be based on
ıpters 1 through 4 and 6 through
be based on Chapters 1 through
ın introduction to epidemiologic
more advanced course detailing
ıgh 5, 9, 10, and 12 through 21.
er collections, depending on the

se they pertain to more than one
ıf the book that relate to a given
that the index be consulted by

ıers, students, and practitioners
ible, but we recognize that any
re grateful to readers of earlier
ıd to continue our past practice
ɔrporating such corrections into
nEpidemiology> to find the latest

sections of the current text and
ıe who helped in that regard, we
Gradus, Jennifer Hill, Katherine
ldani, Henrik Toft Sørensen, Soe
ɔter 18 appeared in the *Interna-*
with permission of Oxford Uni-
er, Gary Friedman, Ulrik Gerdes,
ɔrmation used in the examples of

Kenneth J. Rothman
Sander Greenland
Timothy L. Lash

Contributors

James W. Buehler
Research Professor
Department of Epidemiology
Rollins School of Public Health
Emory University
Atlanta, Georgia

Jack Cahill
Vice President
Department of Health Studies Sector
Westat, Inc.
Rockville, Maryland

Sander Greenland
Professor of Epidemiology and Statistics
University of California
Los Angeles, California

M. Maria Glymour
Robert Wood Johnson Foundation Health and Society Scholar
Department of Epidemiology
Mailman School of Public Health
Columbia University
New York, New York
Department of Society, Human Development and Health
Harvard School of Public Health
Boston, Massachusetts

Marta Gwinn
Associate Director
Department of Epidemiology
National Office of Public Health Genomics
Centers for Disease Control and Prevention
Atlanta, Georgia

Patricia Hartge
Deputy Director
Department of Epidemiology and Biostatistics Program
Division of Cancer Epidemiology and Genetics
National Cancer Institute, National Institutes of Health
Rockville, Maryland

Irva Hertz-Picciotto
Professor
Department of Public Health
University of California, Davis
Davis, California

C. Robert Horsburgh, Jr.
Professor of Epidemiology, Biostatistics and Medicine
Department Epidemiology
Boston University School of Public Health
Boston, Massachusetts

Jay S. Kaufman
Associate Professor
Department of Epidemiology
University of North Carolina at Chapel Hill, School of Public Health
Chapel Hill, North Carolina

Muin J. Khoury
Director
National Office of Public Health Genomics
Centers for Disease Control and Prevention
Atlanta, Georgia

Timothy L. Lash
Associate Professor of Epidemiology and Medicine
Boston University
Boston, Massachusetts

Barbara E. Mahon
Assistant Professor
Department of Epidemiology and Pediatrics
Boston University
Novartis Vaccines and Diagnostics
Boston, Massachusetts

Robert C. Millikan
Professor
Department of Epidemiology
University of North Carolina at Chapel Hill,
School of Public Health
Chapel Hill, North Carolina

Hal Morgenstern
Professor and Chair
Department of Epidemiology
University of Michigan School of
Public Health
Ann Arbor, Michigan

Jørn Olsen
Professor and Chair
Department of Epidemiology
UCLA School of Public Health
Los Angeles, California

Keith O'Rourke
Visiting Assistant Professor
Department of Statistical Science
Duke University
Durham, North Carolina
Adjunct Professor
Department of Epidemiology and
Community Medicine
University of Ottawa
Ottawa, Ontario
Canada

Charles Poole
Associate Professor
Department of Epidemiology
University of North Carolina at Chapel Hill,
School of Public Health
Chapel Hill, North Carolina

Kenneth J. Rothman
Vice President, Epidemiology Research
RTI Health Solutions
Professor of Epidemiology and Medicine
Boston University
Boston, Massachusetts

Clarice R. Weinberg
National Institute of Environmental
Health Sciences
Biostatistics Branch
Research Triangle Park, North Carolina

Noel S. Weiss
Professor
Department of Epidemiology
University of Washington
Seattle, Washington

Allen J. Wilcox
Senior Investigator
Epidemiology Branch
National Institute of Environmental
Health Sciences/NIH
Durham, North Carolina

Walter C. Willett
Professor and Chair
Department of Nutrition
Harvard School of Public Health
Boston, Massachusetts

these hypotheses are not "proved" with the degree of absolute certainty that accompanies the proof of a mathematical theorem.

CAUSAL INFERENCE IN EPIDEMIOLOGY

Etiologic knowledge about epidemiologic hypotheses is often scant, making the hypotheses themselves at times little more than vague statements of causal association between exposure and disease, such as "smoking causes cardiovascular disease." These vague hypotheses have only vague consequences that can be difficult to test. To cope with this vagueness, epidemiologists usually focus on testing the negation of the causal hypothesis, that is, the null hypothesis that the exposure does *not* have a causal relation to disease. Then, any observed association can potentially refute the hypothesis, subject to the assumption (auxiliary hypothesis) that biases and chance fluctuations are not solely responsible for the observation.

TESTS OF COMPETING EPIDEMIOLOGIC THEORIES

If the causal mechanism is stated specifically enough, epidemiologic observations can provide crucial tests of competing, non-null causal hypotheses. For example, when toxic-shock syndrome was first studied, there were two competing hypotheses about the causal agent. Under one hypothesis, it was a chemical in the tampon, so that women using tampons were exposed to the agent directly from the tampon. Under the other hypothesis, the tampon acted as a culture medium for staphylococci that produced a toxin. Both hypotheses explained the relation of toxic-shock occurrence to tampon use. The two hypotheses, however, led to opposite predictions about the relation between the frequency of changing tampons and the rate of toxic shock. Under the hypothesis of a chemical agent, more frequent changing of the tampon would lead to more exposure to the agent and possible absorption of a greater overall dose. This hypothesis predicted that women who changed tampons more frequently would have a higher rate than women who changed tampons infrequently. The culture-medium hypothesis predicts that women who change tampons frequently would have a lower rate than those who change tampons less frequently, because a short duration of use for each tampon would prevent the staphylococci from multiplying enough to produce a damaging dose of toxin. Thus, epidemiologic research, by showing that infrequent changing of tampons was associated with a higher rate of toxic shock, refuted the chemical theory in the form presented. There was, however, a third hypothesis that a chemical in some tampons (e.g., oxygen content) improved their performance as culture media. This chemical-promotor hypothesis made the same prediction about the association with frequency of changing tampons as the microbial toxin hypothesis (Lanes and Rothman, 1990).

Another example of a theory that can be easily tested by epidemiologic data relates to the observation that women who took replacement estrogen therapy had a considerably elevated rate of endometrial cancer. Horwitz and Feinstein (1978) conjectured a competing theory to explain the association: They proposed that women taking estrogen experienced symptoms such as bleeding that induced them to consult a physician. The resulting diagnostic workup led to the detection of endometrial cancer at an earlier stage in these women, as compared with women who were not taking estrogens. Horwitz and Feinstein argued that the association arose from this detection bias, claiming that without the bleeding-induced workup, many of these cancers would not have been detected at all. Many epidemiologic observations were used to evaluate these competing hypotheses. The detection-bias theory predicted that women who had used estrogens for only a short time would have the greatest elevation in their rate, as the symptoms related to estrogen use that led to the medical consultation tended to appear soon after use began. Because the association of recent estrogen use and endometrial cancer was the same in both long- and short-term estrogen users, the detection-bias theory was refuted as an explanation for all but a small fraction of endometrial cancer cases occurring after estrogen use. Refutation of the detection-bias theory also depended on many other observations. Especially important was the theory's implication that there must be a huge reservoir of undetected endometrial cancer in the typical population of women to account for the much greater rate observed in estrogen users, an implication that was not borne out by further observations (Hutchison and Rothman, 1978).

The endometrial cancer example illustrates a critical point in understanding the process of causal inference in epidemiologic studies: Many of the hypotheses being evaluated in the interpretation of epidemiologic studies are auxiliary hypotheses in the sense that they are independent of the presence, absence, or direction of any causal connection between the study exposure and the disease. For example, explanations of how specific types of bias could have distorted an association between exposure and disease are the usual alternatives to the primary study hypothesis. Much of the interpretation of epidemiologic studies amounts to the testing of such auxiliary explanations for observed associations.

CAUSAL CRITERIA

In practice, how do epidemiologists separate causal from noncausal explanations? Despite philosophic criticisms of inductive inference, inductively oriented considerations are often used as criteria for making such inferences (Weed and Gorelic, 1996). If a set of necessary and sufficient causal criteria could be used to distinguish causal from noncausal relations in epidemiologic studies, the job of the scientist would be eased considerably. With such criteria, all the concerns about the logic or lack thereof in causal inference could be subsumed: It would only be necessary to consult the checklist of criteria to see if a relation were causal. We know from the philosophy reviewed earlier that a set of sufficient criteria does not exist. Nevertheless, lists of causal criteria have become popular, possibly because they seem to provide a road map through complicated territory, and perhaps because they suggest hypotheses to be evaluated in a given problem.

A commonly used set of criteria was based on a list of considerations or "viewpoints" proposed by Sir Austin Bradford Hill (1965). Hill's list was an expansion of a list offered previously in the landmark U.S. Surgeon General's report *Smoking and Health* (1964), which in turn was anticipated by the inductive canons of John Stuart Mill (1862) and the rules given by Hume (1739). Subsequently, others, especially Susser, have further developed causal considerations (Kaufman and Poole, 2000).

Hill suggested that the following considerations in attempting to distinguish causal from noncausal associations that were already "perfectly clear-cut and beyond what we would care to attribute to the play of chance": (1) strength, (2) consistency, (3) specificity, (4) temporality, (5) biologic gradient, (6) plausibility, (7) coherence, (8) experimental evidence, and (9) analogy. Hill emphasized that causal inferences cannot be based on a set of rules, condemned emphasis on statistical significance testing, and recognized the importance of many other factors in decision making (Phillips and Goodman, 2004). Nonetheless, the misguided but popular view that his considerations should be used as criteria for causal inference makes it necessary to examine them in detail.

Strength

Hill argued that strong associations are particularly compelling because, for weaker associations, it is "easier" to imagine what today we would call an unmeasured confounder that might be responsible for the association. Several years earlier, Cornfield et al. (1959) drew similar conclusions. They concentrated on a single hypothetical confounder that, by itself, would explain entirely an observed association. They expressed a strong preference for ratio measures of strength, as opposed to difference measures, and focused on how the observed estimate of a risk ratio provides a minimum for the association that a completely explanatory confounder must have with the exposure (rather than a minimum for the confounder–disease association). Of special importance, Cornfield et al. acknowledged that having only a weak association does not rule out a causal connection (Rothman and Poole, 1988). Today, some associations, such as those between smoking and cardiovascular disease or between environmental tobacco smoke and lung cancer, are accepted by most as causal even though the associations are considered weak.

Counterexamples of strong but noncausal associations are also not hard to find; any study with strong confounding illustrates the phenomenon. For example, consider the strong relation between Down syndrome and birth rank, which is confounded by the relation between Down syndrome and maternal age. Of course, once the confounding factor is identified, the association is diminished by controlling for the factor.

These examples remind us that a strong association is neither necessary nor sufficient for causality, and that weakness is neither necessary nor sufficient for absence of causality. A strong association

bears only on hypotheses that the association is entirely or partially due to unmeasured confounders or other source of modest bias.

Consistency

To most observers, consistency refers to the repeated observation of an association in different populations under different circumstances. Lack of consistency, however, does not rule out a causal association, because some effects are produced by their causes only under unusual circumstances. More precisely, the effect of a causal agent cannot occur unless the complementary component causes act or have already acted to complete a sufficient cause. These conditions will not always be met. Thus, transfusions can cause infection with the human immunodeficiency virus, but they do not always do so: The virus must also be present. Tampon use can cause toxic-shock syndrome, but only rarely, when certain other, perhaps unknown, conditions are met. Consistency is apparent only after all the relevant details of a causal mechanism are understood, which is to say very seldom. Furthermore, even studies of exactly the same phenomena can be expected to yield different results simply because they differ in their methods and random errors. Consistency serves only to rule out hypotheses that the association is attributable to some factor that varies across studies.

One mistake in implementing the consistency criterion is so common that it deserves special mention. It is sometimes claimed that a literature or set of results is inconsistent simply because some results are "statistically significant" and some are not. This sort of evaluation is completely fallacious even if one accepts the use of significance testing methods. The results (effect estimates) from a set of studies could all be identical even if many were significant and many were not, the difference in significance arising solely because of differences in the standard errors or sizes of the studies. Conversely, the results could be significantly in conflict even if all were all were nonsignificant individually, simply because in aggregate an effect could be apparent in some subgroups but not others (see Chapter 33). The fallacy of judging consistency by comparing P-values or statistical significance is not eliminated by "standardizing" estimates (i.e., dividing them by the standard deviation of the outcome, multiplying them by the standard deviation of the exposure, or both); in fact it is worsened, as such standardization can create differences where none exists, or mask true differences (Greenland et al., 1986, 1991; see Chapters 21 and 33).

Specificity

The criterion of specificity has two variants. One is that a cause leads to a single effect, not multiple effects. The other is that an effect has one cause, not multiple causes. Hill mentioned both of them. The former criterion, specificity of effects, was used as an argument in favor of a causal interpretation of the association between smoking and lung cancer and, in an act of circular reasoning, in favor of ratio comparisons and not differences as the appropriate measures of strength. When ratio measures were examined, the association of smoking to diseases looked "quantitatively specific" to lung cancer. When difference measures were examined, the association appeared to be nonspecific, with several diseases (other cancers, coronary heart disease, etc.) being at least as strongly associated with smoking as lung cancer was. Today we know that smoking affects the risk of many diseases and that the differenee comparisons were accurately portraying this lack of specificity. Unfortunately, however, the historical episode of the debate over smoking and health is often cited today as justification for the specificity criterion and for using ratio comparisons to measure strength of association. The proper lessons to learn from that episode should be just the opposite.

Weiss (2002) argued that specificity can be used to distinguish some causal hypotheses from noncausal hypotheses, when the causal hypothesis predicts a relation with one outcome but no relation with another outcome. His argument is persuasive when, in addition to the causal hypothesis, one has an alternative noncausal hypothesis that predicts a nonspecific association. Weiss offered the example of screening sigmoidoscopy, which was associated in case-control studies with a 50% to 70% reduction in mortality from distal tumors of the rectum and tumors of the distal colon, within the reach of the sigmoidoscope, but no reduction in mortality from tumors elsewhere in the colon. If the effect of screening sigmoidoscopy were not specific to the distal colon tumors, it would lend support not to all noncausal theories to explain the association, as Weiss suggested, but only to those noncausal theories that would have predicted a nonspecific association. Thus, specificity can

come into play when it can be logically deduced from the causal hypothesis in question and when nonspecificity can be logically deduced from one or more noncausal hypotheses.

Temporality

Temporality refers to the necessity that the cause precede the effect in time. This criterion is inarguable, insofar as any claimed observation of causation must involve the putative cause C preceding the putative effect D. It does *not*, however, follow that a reverse time order is evidence against the hypothesis that C can cause D. Rather, observations in which C followed D merely show that C could not have caused D in these instances; they provide no evidence for or against the hypothesis that C can cause D in those instances in which it precedes D. Only if it is found that C cannot precede D can we dispense with the causal hypothesis that C *could* cause D.

Biologic Gradient

Biologic gradient refers to the presence of a dose–response or exposure–response curve with an expected shape. Although Hill referred to a "linear" gradient, without specifying the scale, a linear gradient on one scale, such as the risk, can be distinctly nonlinear on another scale, such as the log risk, the odds, or the log odds. We might relax the expectation from linear to strictly monotonic (steadily increasing or decreasing) or even further merely to monotonic (a gradient that never changes direction). For example, more smoking means more carcinogen exposure and more tissue damage, hence more opportunity for carcinogenesis. Some causal associations, however, show a rapid increase in response (an approximate threshold effect) rather than a strictly monotonic trend. An example is the association between DES and adenocarcinoma of the vagina. A possible explanation is that the doses of DES that were administered were all sufficiently great to produce the maximum effect from DES. Under this hypothesis, for all those exposed to DES, the development of disease would depend entirely on other component causes.

The somewhat controversial topic of alcohol consumption and mortality is another example. Death rates are higher among nondrinkers than among moderate drinkers, but they ascend to the highest levels for heavy drinkers. There is considerable debate about which parts of the J-shaped dose–response curve are causally related to alcohol consumption and which parts are noncausal artifacts stemming from confounding or other biases. Some studies appear to find only an increasing relation between alcohol consumption and mortality, possibly because the categories of alcohol consumption are too broad to distinguish different rates among moderate drinkers and nondrinkers, or possibly because they have less confounding at the lower end of the consumption scale.

Associations that do show a monotonic trend in disease frequency with increasing levels of exposure are not necessarily causal. Confounding can result in a monotonic relation between a noncausal risk factor and disease if the confounding factor itself demonstrates a biologic gradient in its relation with disease. The relation between birth rank and Down syndrome mentioned earlier shows a strong biologic gradient that merely reflects the progressive relation between maternal age and occurrence of Down syndrome.

These issues imply that the existence of a monotonic association is neither necessary nor sufficient for a causal relation. A nonmonotonic relation only refutes those causal hypotheses specific enough to predict a monotonic dose–response curve.

Plausibility

Plausibility refers to the scientific plausibility of an association. More than any other criterion, this one shows how narrowly systems of causal criteria are focused on epidemiology. The starting point is an epidemiologic association. In asking whether it is causal or not, one of the considerations we take into account is its plausibility. From a less parochial perspective, the entire enterprise of causal inference would be viewed as the act of determining how plausible a causal *hypothesis* is. One of the considerations we would take into account would be epidemiologic associations, if they are available. Often they are not, but causal inference must be done nevertheless, with inputs from toxicology, pharmacology, basic biology, and other sciences.

Just as epidemiology is not essential for causal inference, plausibility can change with the times. Sartwell (1960) emphasized this point, citing remarks of Cheever in 1861, who had been commenting on the etiology of typhus before its mode of transmission (via body lice) was known:

> It could be no more ridiculous for the stranger who passed the night in the steerage of an emigrant ship to ascribe the typhus, which he there contracted, to the vermin with which bodies of the sick might be infested. An adequate cause, one reasonable in itself, must correct the coincidences of simple experience.

What was to Cheever an implausible explanation turned out to be the correct explanation, because it was indeed the vermin that caused the typhus infection. Such is the problem with plausibility: It is too often based not on logic or data, but only on prior beliefs. This is not to say that biologic knowledge should be discounted when a new hypothesis is being evaluated, but only to point out the difficulty in applying that knowledge.

The Bayesian approach to inference attempts to deal with this problem by requiring that one quantify, on a probability (0 to 1) scale, the certainty that one has in prior beliefs, as well as in new hypotheses. This quantification displays the dogmatism or open-mindedness of the analyst in a public fashion, with certainty values near 1 or 0 betraying a strong commitment of the analyst for or against a hypothesis. It can also provide a means of testing those quantified beliefs against new evidence (Howson and Urbach, 1993). Nevertheless, no approach can transform plausibility into an objective causal criterion.

Coherence

Taken from the U.S. Surgeon General's *Smoking and Health* (1964), the term *coherence* implies that a cause-and-effect interpretation for an association does not conflict with what is known of the natural history and biology of the disease. The examples Hill gave for coherence, such as the histopathologic effect of smoking on bronchial epithelium (in reference to the association between smoking and lung cancer) or the difference in lung cancer incidence by sex, could reasonably be considered examples of plausibility, as well as coherence; the distinction appears to be a fine one. Hill emphasized that the absence of coherent information, as distinguished, apparently, from the presence of conflicting information, should not be taken as evidence against an association being considered causal. On the other hand, the presence of conflicting information may indeed refute a hypothesis, but one must always remember that the conflicting information may be mistaken or misinterpreted. An example mentioned earlier is the "inhalation anomaly" in smoking and lung cancer, the fact that the excess of lung cancers seen among smokers seemed to be concentrated at sites in the upper airways of the lung. Several observers interpreted this anomaly as evidence that cigarettes were not responsible for the excess. Other observations, however, suggested that cigarette-borne carcinogens were deposited preferentially where the excess was observed, and so the anomaly was in fact consistent with a causal role for cigarettes (Wald, 1985).

Experimental Evidence

To different observers, experimental evidence can refer to clinical trials, to laboratory experiments with rodents or other nonhuman organisms, or to both. Evidence from human experiments, however, is seldom available for epidemiologic research questions, and animal evidence relates to different species and usually to levels of exposure very different from those that humans experience. Uncertainty in extrapolations from animals to humans often dominates the uncertainty of quantitative risk assessments (Freedman and Zeisel, 1988; Crouch et al., 1997).

To Hill, however, experimental evidence meant something else: the "experimental, or semi-experimental evidence" obtained from reducing or eliminating a putatively harmful exposure and seeing if the frequency of disease subsequently declines. He called this the strongest possible evidence of causality that can be obtained. It can be faulty, however, as the "semi-experimental" approach is nothing more than a "before-and-after" time trend analysis, which can be confounded or otherwise biased by a host of concomitant secular changes. Moreover, even if the removal of exposure does causally reduce the frequency of disease, it might not be for the etiologic reason hypothesized. The draining of a swamp near a city, for instance, would predictably and causally reduce the rate of yellow fever or malaria in that city the following summer. But it would be a mistake to call this observation the strongest possible evidence of a causal role of miasmas (Poole, 1999).

Analogy

Whatever insight might be derived from analogy is handicapped by the inventive imagination of scientists who can find analogies everywhere. At best, analogy provides a source of more elaborate hypotheses about the associations under study; absence of such analogies reflects only lack of imagination or experience, not falsity of the hypothesis.

We might find naive Hill's examples in which reasoning by analogy from the thalidomide and rubella tragedies made it more likely to him that other medicines and infections might cause other birth defects. But such reasoning is common; we suspect most people find it more credible that smoking might cause, say, stomach cancer, because of its associations, some widely accepted as causal, with cancers in other internal and gastrointestinal organs. Here we see how the analogy criterion can be at odds with either of the two specificity criteria. The more apt the analogy, the less specific are the effects of a cause or the less specific the causes of an effect.

Summary

As is evident, the standards of epidemiologic evidence offered by Hill are saddled with reservations and exceptions. Hill himself was ambivalent about their utility. He did not use the word *criteria* in the speech. He called them "viewpoints" or "perspectives." On the one hand, he asked, "In what circumstances can we pass from this observed *association* to a verdict of *causation*?" (emphasis in original). Yet, despite speaking of verdicts on causation, he disagreed that any "hard-and-fast rules of evidence" existed by which to judge causation: "None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non" (Hill, 1965).

Actually, as noted above, the fourth viewpoint, temporality, is *a sine qua non* for causal explanations of observed associations. Nonetheless, it does not bear on the hypothesis that an exposure is capable of causing a disease in situations as yet unobserved (whether in the past or the future). For suppose every exposed case of disease ever reported had received the exposure after developing the disease. This reversed temporal relation would imply that exposure had not caused disease among these reported cases, and thus would refute the hypothesis that it had. Nonetheless, it would *not* refute the hypothesis that the exposure is *capable* of causing the disease, or that it had caused the disease in unobserved cases. It would mean only that we have no worthwhile epidemiologic evidence relevant to that hypothesis, for we had not yet seen what became of those exposed before disease occurred relative to those unexposed. Furthermore, what appears to be a causal sequence could represent reverse causation if preclinical symptoms of the disease lead to exposure, and then overt disease follows, as when patients in pain take analgesics, which may be the result of disease that is later diagnosed, rather than a cause.

Other than temporality, there is no necessary or sufficient criterion for determining whether an observed association is causal. Only when a causal hypothesis is elaborated to the extent that one can predict from it a particular form of consistency, specificity, biologic gradient, and so forth, can "causal criteria" come into play in evaluating causal hypotheses, and even then they do not come into play in evaluating the general hypothesis per se, but only some specific causal hypotheses, leaving others untested.

This conclusion accords with the views of Hume and many others that causal inferences cannot attain the certainty of logical deductions. Although some scientists continue to develop causal considerations as aids to inference (Susser, 1991), others argue that it is detrimental to cloud the inferential process by considering checklist criteria (Lanes and Poole, 1984). An intermediate, refutationist approach seeks to transform proposed criteria into deductive tests of causal hypotheses (Maclure, 1985; Weed, 1986). Such an approach helps avoid the temptation to use causal criteria simply to buttress pet theories at hand, and instead allows epidemiologists to focus on evaluating competing causal theories using crucial observations. Although this refutationist approach to causal inference may seem at odds with the common implementation of Hill's viewpoints, it actually seeks to answer the fundamental question posed by Hill, and the ultimate purpose of the viewpoints he promulgated:

> What [the nine viewpoints] can do, with greater or less strength, is to help us to make up our minds on the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect? (Hill, 1965)

The crucial phrase "equally or more likely than cause and effect" suggests to us a subjective assessment of the certainty, or probability of the causal hypothesis at issue relative to another hypothesis. Although Hill wrote at a time when expressing uncertainty as a probability was unpopular in statistics, it appears from his statement that, for him, causal inference is a subjective matter of degree of personal belief, certainty, or conviction. In any event, this view is precisely that of subjective Bayesian statistics (Chapter 18).

It is unsurprising that case studies (e.g., Weed and Gorelick, 1996) and surveys of epidemiologists (Holman et al., 2001) show, contrary to the rhetoric that often attends invocations of causal criteria, that epidemiologists have *not* agreed on a set of causal criteria or on how to apply them. In one study in which epidemiologists were asked to employ causal criteria to fictional summaries of epidemiologic literatures, the agreement was only slightly greater than would have been expected by chance (Holman et al., 2001). The typical use of causal criteria is to make a case for a position for or against causality that has been arrived at by other, unstated means. Authors pick and choose among the criteria they deploy, and define and weight them in *ad hoc* ways that depend only on the exigencies of the discussion at hand. In this sense, causal criteria appear to function less like standards or principles and more like values (Poole, 2001b), which vary across individual scientists and even vary within the work of a single scientist, depending on the context and time. Thus universal and objective causal criteria, if they exist, have yet to be identified.

Exhibit 117

EPIDEMIOLOGY

Concepts and Methods

William A. Oleckno
Northern Illinois University

WAVELAND PRESS, INC.
Long Grove, Illinois

CHAPTER SEVEN

Association and Causation in Epidemiology

This chapter discusses differences among spurious, noncausal, and causal associations, the various types of causes, and common guidelines used in assessing causation in epidemiologic studies.

Learning Objectives

- Describe and give examples of spurious, noncausal, and causal associations in epidemiology.
- State the common reasons for spurious and noncausal associations, respectively.
- Distinguish among necessary, sufficient, necessary and sufficient, necessary but not sufficient, not necessary but sufficient, and not necessary and not sufficient causes and give examples of each type.
- Describe and give examples of direct and indirect causal associations.
- Briefly describe the causal pie model.
- Discuss six guidelines based on Hill's postulates for judging potential causal associations, including the advantages and limitations of each criterion, respectively.
- Explain the importance of finding causal associations in epidemiology.
- Define predisposing or enabling factors, statistical association, and threshold.

Introduction

As indicated in chapter 1, one of the primary goals of epidemiology is to discover the *causes** of morbidity and mortality in human populations. This goal has immense practical significance for health professionals because a better

*There are many terms relating to or derived from the root term *cause*. These include causation, causality, causal, causative, cause-effect, etiology, and so forth. These terms are not defined separately in this chapter, but each refers to something similar.

tion likely due to random error, bias, confounding, or a reserved causal sequence? This may take some critical thinking, further analysis, or consultation. If these seem to be unlikely explanations, it can be helpful to review some generally accepted guidelines for establishing causation such as those described by Sir Austin Bradford Hill.

In 1965, Sir Austin Bradford Hill, Professor Emeritus of Medical Statistics with the University of London, delivered a landmark address where he outlined nine criteria that could be used to determine if statistical associations were likely to represent causal associations.[11] His reasoning built on the earlier work of others, such as John Stuart Mill, who in 1856 had defined several canons from which causal relationships could be deduced.[6] Over the years many authors have articulated or modified Hill's basic criteria, which have become known as **Hill's postulates**. Using these as a focal point, the following six guidelines should be helpful in deciding whether or not statistical associations are likely to represent causal associations (figure 7-1). In the end, the process of determining causation is largely subjective except for the first guideline, which is actually a requirement.

- **Correct temporal sequence.** In order for an exposure to be considered a cause of an outcome, it must *precede* the outcome. Of all the guidelines used to judge whether an association is causal or not, this is the only one that is considered *absolutely essential*. Exposures that occur concurrently with an outcome or subsequent to an outcome cannot be considered causal because they do not alter the frequency of the outcome. Determining if an exposure precedes an outcome can be problematic in cross-sectional studies where exposure and outcome are assessed concurrently. For example, in a cross-sectional study designed to determine if there is a relationship between the prevalence of excess body weight and osteoarthritis, it may not be clear which factor came first. Thus, the correct temporal sequence cannot be established reliably. This can also be a problem in case-control studies where the prevalence of the outcome is assessed instead of its incidence.
- **Strength of the association.** In general, the stronger an association between a given exposure and outcome (see table 6-3), the more likely the association is causal. When the risk ratio is very high, for example, it is more difficult to explain away the association due to unrecognized or subtle sources of bias or confounding. Compared to nonsmokers, those who smoke and are exposed to high levels of asbestos in their jobs have a fifty- to ninety-fold increased risk of lung cancer. It seems improbable that these factors are not causative. Even if some bias or confounding exists, it is unlikely that it would account for the entire relationship. This is not to say that small associations cannot also be causal in nature. This is one reason why several guidelines are needed to assess causality.
- **Consistency of the association.** When other investigators studying different populations at different times in different places using different methodologies obtain similar findings with regard to a specific association, it